

Universidade de Lisboa
Faculdade de Medicina de Lisboa



**Análise genética com Y-SNPs de subgrupos populacionais
residentes no Sul de Portugal**

Sara Alexandra Patrício Gaudêncio

Orientadores: **Mestre Cláudia Vieira da Silva**
Prof. Doutor Jorge Costa Santos

**Dissertação especialmente elaborada para obtenção do grau de Mestre em
Medicina Legal e Ciências Forenses**

2016

Universidade de Lisboa
Faculdade de Medicina de Lisboa



**Análise genética com Y-SNPs de subgrupos populacionais
residentes no Sul de Portugal**

Sara Alexandra Patrício Gaudêncio

Orientadores: **Mestre Cláudia Vieira da Silva**
Prof. Doutor Jorge Costa Santos

**Dissertação especialmente elaborada para obtenção do grau de Mestre em
Medicina Legal e Ciências Forenses**

2016

Todas as afirmações efetuadas no presente documento são da exclusiva responsabilidade do seu autor, não cabendo qualquer responsabilidade à Faculdade de Medicina de Lisboa pelos conteúdos nele apresentados.

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 19 de janeiro de 2016.

AGRADECIMENTOS

A realização deste mestrado assim como a respetiva dissertação não teria sido possível sem a ajuda de todas as pessoas que me apoiaram em mais uma etapa da minha académica.

A todos eles os meus agradecimentos, por terem contribuído para a concretização dos meus objetivos. Em particular:

Ao Professor Doutor Jorge Costa Santos, por ter aceitado o meu projeto, dando-me a possibilidade de realizar a dissertação de mestrado no Instituto de Medicina Legal e Ciências Forenses – Delegação Sul. Por ter acedido ser meu coorientador, por todo o conhecimento científico, pelas valiosas sugestões e por toda a amabilidade.

À Diretora do Serviço de Biologia e Genética Forense do INMLCF – Sul, Dr.^a Teresa Ribeiro e à Diretora Nacional do Serviço de Biologia e Genética Forense do INMLCF, Dr.^a Maria Teresa Porto por me terem recebido e permitido realizar o trabalho laboratorial neste Serviço.

À minha orientadora, Mestre Cláudia Vieira da Silva, por ter anuído ser orientadora desta dissertação. Por toda a disponibilidade, pelos conhecimentos transmitidos, pela paciência, por ter sempre uma palavra amiga e um sorriso na cara mesmo quando os resultados teimavam em não aparecer.

A todas as pessoas do Serviço, nomeadamente à Mestre Heloísa Costa e ao Mestre António Amorim por toda a ajuda, pelos conselhos e pelos conhecimentos partilhados. À D^a Conceição, por toda a boa disposição e entusiasmo que traz ao serviço e que contagiou até quando sentia que estava a fraquejar.

Às minhas colegas de laboratório, Joana e Filipa, pelos conhecimentos laboratoriais transmitidos, sem dúvida foram imprescindíveis. Às minhas colegas de mestrado, Ana e Marisa, pelas gargalhadas, brincadeiras e por toda a entajuda para a concretização desta dissertação.

Aos meus pais, Luís e Lúcia, por me terem permitido realizar este mestrado e pelos sacrifícios feitos para o concretizar dos meus sonhos académicos. À minha irmã, Joana, por ser minha parceira dos devaneios quando o desespero e o cansaço começavam a dominar.

Ao meu namorado, Mauro, por toda a amizade e paciência. Obrigado por ouvires os meus desabafos, pelas palavras de incentivo e sempre com um sorriso guardado para me animar.

Por fim, aos meus amigos, pela compreensão nos momentos em que estive mais ausente. Por nunca se esquecerem de mim ao longo dos meus anos de vida académica.

A todos, mais uma vez, os meus sinceros agradecimentos.

ÍNDICE

AGRADECIMENTOS	4
ÍNDICE	6
RESUMO.....	9
ABSTRACT	11
INTRODUÇÃO	13
1.1 Genoma Humano	13
1.1.1 Marcadores genéticos.....	15
1.1.2 Mutação Genética.....	16
1.2 Cromossoma Y.....	17
1.3 SNPs	20
1.3.1 SNPs: Aplicações Forenses	23
1.3.2 SNPs do Cromossoma Y: Aplicações Forenses	24
1.4 SNPs vs STRs.....	25
1.5 Árvore dos Haplogrupos Binários do Cromossoma Y	28
1.6 Populações em estudo	32
1.6.1 Cabo Verde.....	32
1.6.2 Portugal	36
1.7 Objetivos do Estudo	39
MATERIAL E MÉTODOS	40

2.1	Colheita e Armazenamento de Amostra	40
2.2	Extração de ADN	41
2.3	Quantificação de ADN	43
2.4	Genotipagem de Y- SNPs	44
2.4.1	Amplificação de ADN	47
2.4.2	Purificação com <i>ExoSAP-IT</i> ®	51
2.4.2	Minisequenciação com SNaPShot™ Multiplex Kit	53
2.4.3	Purificação com SAP	58
2.5	Deteção e Análise do Produto Minisequenciado	59
2.5.1	Deteção do Produto Sequenciado	61
2.6	Análise estatística, designação de Haplogrupos e Árvore Filogenética	62
RESULTADOS E DISCUSSÃO		64
3.1	Análise do Multiplex: Estudo dos Polimorfismos	67
3.2	Cálculo das frequências: Estudo dos Haplogrupos	69
3.3	Distância Genética	74
CONCLUSÕES		76
REFERÊNCIAS BIBLIOGRÁFICAS		78
ANEXOS.....		90
Anexo I – Protocolo utilizado para Extração de ADN a partir de manchas de sangue por Chelex® 100		90

Anexo II – Protocolo utilizado para Quantificação de ADN com o Quantifiler® Human ADN Quantification Kit	92
Anexo III – Protocolo utilizado para Amplificação de ADN por PCR com o QIAGEN® Multiplex PCR Kit.....	94
Anexo IV – Protocolo utilizado para Minisequenciação com SNaPshot™ Multiplex Kit.....	96
Anexo V – Protocolo utilizado para Eletroforese Capilar no sequenciador automático <i>ABI Prism® 3130 GeneticAnalyzer</i>	98
Anexo VI – Tabela de Resultados.....	99
ÍNDICE DE FIGURAS	105
ÍNDICE DE TABELAS	108
ÍNDICE DE ABREVIATURAS.....	110

RESUMO

Na genética forense, o estudo do cromossoma Y é de particular interesse devido às suas características únicas, nomeadamente a não recombinação meiótica. Como tal, a informação genética é transmitida integralmente de pai para filho. A não existência de alterações genéticas, salvo mutações esporádicas, possibilita a análise da linhagem paterna, permitindo estudos de ancestralidade e dos fenómenos migratórios.

A identificação das linhagens é realizada com marcadores bialélicos, SNPs (*Single Nucleotide Polymorphisms*), do cromossoma Y. Consistem numa única mutação em determinada posição genómica do cromossoma, presente em mais de 1% da população, permitindo determinar a origem geográfica de um indivíduo. São de particular importância pois têm uma baixa taxa de mutação, são abundantes no genoma e permitem a análise de amostras de ADN degradadas.

Cabo Verde é um país já com longa relação histórica com Portugal, proveniente do século XV, com a descoberta deste arquipélago e a sua colonização. Atualmente, são a segunda maior comunidade de imigrantes em Portugal, com cerca de 42.301 cabo-verdianos, residentes principalmente nas regiões de Lisboa, Setúbal e Algarve. Para este trabalho de investigação, analisou-se amostras de 75 caucasianos com origem no sul de Portugal e 75 indivíduos naturais de Cabo Verde que compareceram no Serviço de Biologia e Genética Forense da Delegação Sul do INMLCF, IP. no âmbito de investigações de parentesco.

O estudo dos Y-SNPs foi realizado segundo a metodologia de *Brión et al.* (2004), com um *multiplex* constituído por 9 marcadores: 92R7, M70, M22, Tat, P25,

SRY1532, M173, M213 e M9. Foram minisequenciados com o *SNaPShot™ Multiplex Kit*.

Depois de obtidos os resultados, foram caracterizados os haplogrupos segundo a Árvore dos Haplogrupos Binários do Cromossoma Y, proposta pelo YCC e revista por *Jobling e Tyler-Smith* (2003) e por *Karafet et al.* (2008). Foram calculadas as suas frequências e a distância genética entre estes dois grupos populacionais.

Palavras-chave: Genética forense, Cromossoma Y, SNPs, Haplogrupos, Cabo Verde.

ABSTRACT

In forensic genetics, the study of Y chromosome is of particular interest due to their unique characteristics, including non-recombination, as such, the genetic information is transmitted from father to son integrally. The absence of genetic alterations, except sporadic mutations, enables the analysis of the paternal lineage, allowing studies of ancestry and migration.

The identification of the strains is carried out with biallelic markers of the Y - chromosome, SNPs. They consist of a single mutation in a given genomic position of the chromosome, present in more than 1% of the population, allowing determine the geographical origin of an individual. They have particular importance since they have a low mutation rate, are abundant in the genome, and allow the analysis of degraded DNA samples.

Cape Verde is a country with a long historical relationship with Portugal, from the fifteenth century, with the discovery and colonization of this archipelago. Actually, is the second largest immigrant community in Portugal, with about 42,301 Cape Verdeans who live mainly in the regions of Lisbon, Setúbal and Algarve.

For this research, samples were analyzed for 75 caucasians of Portugal origins and 75 of Cape Verde origins who attended the Serviço de Biologia e Genética Forense da Delegação do Sul do INMLCF, IP. For paternity investigation.

The study of Y-SNPs was carried out according to the methodology described by Brion et al. (2004), with a multiplex constituted by 9 markers: 92R7, M70, M22, Tat, P25, SRY1532, M173, M213 and M9. This method for allele discrimination was used the SNaPShot™ Multiplex Kit.

After the results obtained, they were characterized according to the Tree of Y-Chromosomal Binary Haplogroups, proposed by the YCC and reviewed by Jobling and Tyler-Smith (2003) and Karafet et al. (2008).

The frequency of the haplogroups and the genetic distance between these two population groups were calculated.

Keywords: Forensic Genetics, Y chromosome, SNPs, Haplogroups, Cape Verde.

INTRODUÇÃO

1.1 Genoma Humano

O Genoma Humano é o conjunto de toda a informação genética humana e está contida no interior de todas as células diploides. Esta informação está codificada na molécula designada por Ácido Desoxirribonucleico – ADN (Makalowski, 2001; Martins e Carvalho, 2008) e corresponde ao genoma nuclear, onde está armazenada em 23 pares de cromossomas (22 autossómicos e 1 sexual – XX será mulher, XY será homem) confinados no núcleo das células e ao genoma mitocondrial, presente nas mitocôndrias dispersas pelo citoplasma celular (Brisighelli *et al.*, 2006; Goodwin, Linacre e Hadi, 2011; Makalowski, 2001).

A macromolécula ADN consiste em duas cadeias polinucleotídicas, complementares e antiparalelas, dispostas em dupla hélice. É composta por quatro entidades, que por sua vez se subdividem em 3 unidades: açúcar (desoxirribose), uma base azotada e um grupo fosfato (Butler, 2005; Grilo, 2001). As bases azotadas (Timina (T), Adenina (A), Citosina (C) e Guanina (G)) em conjunto com o açúcar e o grupo fosfato formam os nucleótidos. Estes nucleótidos são discriminados pela letra que representa a sua base azotada. As bases azotadas dividem-se em dois grupos: purinas (A e G) e pirimidinas (T e C). A adenina de uma cadeia liga-se exclusivamente à timina da outra, ou seja, cada nucleótido só emparelha com o seu complementar formando pontes de hidrogénio e que se passam a designar por pares de bases (pb) (Fig.1) (Butler, 2010; Grilo, 2001; Martins e Carvalho, 2008; Nussbaum, McInnes e Willard, 2015).

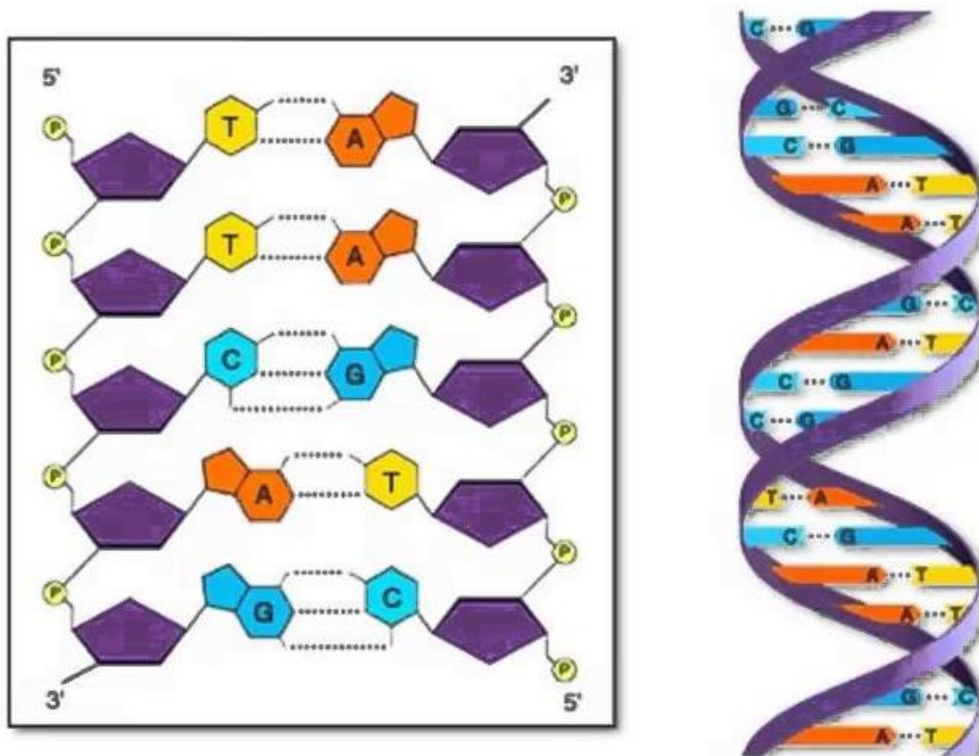


Figura 1 - Representação da constituição de uma molécula de ADN, com o açúcar, o grupo fosfato e a base azotada. O nucleótido tem a letra que representa a base azotada presente. Adaptado de: <http://www.mas-que-ciencia.com/category/genetica/genetica-molecular>.

O ADN nuclear (ADNnu), presente em todas as células encontra-se organizado em cromossomas, altamente compactado, e revestido por histonas que lhe conferem proteção (Pinheiro, 2008). O ADNnu tem como funções principais a realização de cópias de si mesmo, mantendo a informação genética completa e idêntica para que a produção de proteínas não sofra qualquer alteração (Butler, 2005; Goodwin, Linacre e Hadi, 2011; Martins e Carvalho, 2008).

1.1.1 Marcadores genéticos

O material genético, no cromossoma, é constituído por regiões codificantes (correspondentes a 5% do genoma) e regiões não codificantes (95% do genoma) (Butler, 2005, 2010).

As regiões codificantes são constituídas por genes repartidas em exões (sequências que codificam) e intrões (sequências estruturais). Estas regiões que contêm a informação necessária para a produção de proteínas e que revelam, por isso, interesse médico, não sendo utilizado na área forense (Butler, 2005, 2010).

As regiões não codificantes são regiões bastante polimórficas na qual estão localizados os marcadores genéticos com interesse forense. Estas regiões são constituídas por sequências repetitivas de ADN importantes para a identificação de um indivíduo (Butler, 2010; Pinheiro, 2010).

As sequências repetitivas de ADN estão localizadas numa determinada posição cromossómica, designada por *locus*. Em cada *locus* existe uma das formas do marcador genético, que se dá o nome de alelo. Se num determinado *locus*, um indivíduo apresentar alelos idênticos denomina-se por homozigótico. Se apresentar dois alelos distintos no mesmo *locus*, será heterozigótico. Os resultados observados nos alelos de um *locus* designam-se por genótipo e a caracterização dos diversos genótipos obtidos em diversos *loci*, permite a construção de um perfil genético que resulta então da combinação de múltiplos genótipos estudados (Butler, 2005, 2010).

A informação contida na molécula de ADN é transmitida à descendência, de geração em geração. Assim, cada indivíduo contém no seu genoma metade da informação que deriva do genoma da mãe e a restante metade deriva do genoma do pai. No entanto, para serem utilizados para identificação do indivíduo, os

marcadores de ADN deverão ter como atributos o elevado polimorfismo, elevada variabilidade, baixa taxa de mutação e capacidade de discriminação das amostras (Gusmão e Alves, 2010; Nussbaum, McInnes e Willard, 2015; Pinheiro, 2010).

1.1.2 Mutação Genética

As células humanas sofrem divisão de forma contínua, como consequência existe uma maior probabilidade de ocorrência de erros. Todavia, as células têm a capacidade de “reparar” os erros introduzidos no seu ADN. Quando estas alterações se mantêm, mesmo após uma correção, designa-se então por mutação e define-se: “todas e quaisquer alterações na sequência de pares de bases do ADN” (Gusmão e Alves, 2010).

As mutações podem ser pontuais, por substituição, inserção ou deleção, duplicação ou translocação e podem apresentar um carácter vantajoso, desvantajoso ou, como ocorre na sua maioria, um carácter neutro (Nussbaum, McInnes e Willard, 2015).

Quando a mutação sucede, existe uma alteração nos nucleótidos, constituindo-se um novo alelo. Esta variabilidade, quando registada em mais do que 1% da população, transpõe-se para um polimorfismo (Nussbaum, McInnes e Willard, 2015).

Um polimorfismo, que ocorre ao nível da linha germinativa, é bastante suscetível de ser transmitido à descendência, revelando interesse forense, pois pressupõe-se que seja uma variabilidade já estável para que seja realmente transmitida às gerações seguintes e em simultâneo que permita um *locus* informativo (Pinheiro, 2010).

1.2 Cromossoma Y

Os cromossomas sexuais, X e Y, têm origem comum, ou seja, são geneticamente homólogos (Charlesworth, 1991). Porém, a ocorrência de restrições de recombinação genética e mutações que se foram sucedendo ao longo de centenas de milhares de anos resultaram na redução de atividade do cromossoma Y e consequentemente na redução do seu tamanho (Willard, 2003).

O cromossoma Y é, em tamanho, o terceiro cromossoma humano mais pequeno com um tamanho de aproximadamente 60 milhões de pares de bases (Mb), valor bastante inferior ao do seu homólogo cromossoma X, que terá cerca de 155 Mb (Butler, 2005; Gusmão e Alves, 2010).

Este cromossoma, presente em todas as células nucleadas de um indivíduo de género masculino, é de valiosa importância no domínio forense pois a sua transmissão ocorre de forma praticamente inalterada de pai para filho. Assim, ao proceder-se à análise de uma amostra para o estudo do cromossoma Y com fins forenses, simultaneamente procede-se à caracterização de toda a linhagem paterna daquele indivíduo (Jobling, 2001).

O cromossoma masculino possui características muito particulares. É verdadeiramente haploide e é transmitido à geração seguinte como um bloco (Jobling, 2001; Nussbaum, McInnes e Willard, 2015; Serdar e Sema, 2009). Apenas as suas extremidades (regiões autossómicas) denominadas por PAR 1 – situada na extremidade do braço curto (Yp), com cerca de 2,6 Mb e por PAR 2 – presente na extremidade do braço longo (Yq), com um tamanho de 0,32 Mb, sofrem recombinação genética com o cromossoma X durante o processo da meiose (Lahn *et al.*, 2001). A restante estrutura, deste cromossoma é formada pelas regiões da

Heterocromatina e Eucromatina. A primeira está localizada no braço longo (Yq) e é composta por sequências de ADN repetitivo. A segunda é uma região constituída por sequências de ADN repetitivas e genes codificantes de proteínas. É nesta zona eucromática que se encontra o gene masculino SRY, responsável pelo controlo do desenvolvimento sexual masculino (Fig.2) (Butler, 2005, 2010, 2012; Gusmão e Alves, 2010; Li *et al.*, 2006; Wetton, Tsang e Khan, 2005).

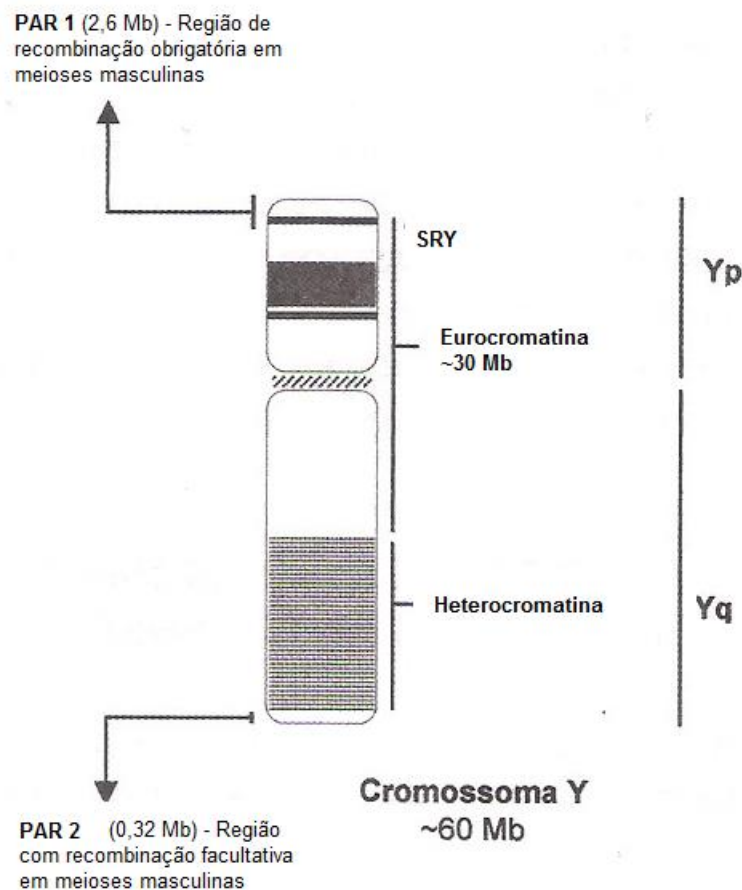


Figura 2 - Cromossoma Y e os seus constituintes. Adaptado de Gusmão e Alves, 2010.

De referir que, aproximadamente 95% do cromossoma é designada como Região Não Recombinante do Cromossoma Y (NRY), ou como é definido atualmente, Região do Cromossoma Y Específica Masculina (MSY) (Bouakaze *et al.*, 2007; Skaletsky *et al.*, 2003; Underhill *et al.*, 2000). Esta deliberação é feita devido ao facto de esta região não sofrer qualquer recombinação durante a meiose (Adams *et al.*, 2006). Assim, o cromossoma Y contém a informação da totalidade de mutações que acumulou no passado, permitindo a construção da linhagem paterna. Conclui-se que a região MSY de todos os cromossomas Y partilha o mesmo ancestral paterno (Gusmão e Alves, 2010; Hughes *et al.*, 2005; Skaletsky *et al.*, 2003).

O perfil obtido baseado nos marcadores presentes neste cromossoma, tem informação relativa a ADN não recombinante, sendo partilhado por vários indivíduos, e portanto, não tem o mesmo poder de discriminação que os marcadores autossómicos (Gusmão e Alves, 2010; Jobling e Tyler-Smith, 2000).

Contudo, os marcadores do cromossoma Y são de extrema utilidade em muitos casos forenses, nomeadamente Investigações de Parentesco complexas em que o filho é do sexo masculino e não é possível a recolha de amostra do presumível pai, em casos específicos de identificação com a amostra em estado avançado de degradação como em acidentes de massa. Por serem específicos do género masculino, assumem especial importância nos casos criminais em que existe mistura de amostras biológicas, do criminoso – género masculino – e da vítima – género feminino.

A aplicação de Y-SNPs pode ser uma ferramenta essencial para estudar a evolução do homem moderno, através da migração das populações (Jobling,

Pandya e Tyler-Smith, 1997; Sanchez, Børsting e Morling, 2005) e nos casos de paternidade complexos, (Butler, 2012; Gusmão e Alves, 2010).

1.3 SNPs

Anteriormente à introdução de ADN na rotina forense, as perícias eram elaboradas com base na caracterização de marcadores sorológicos clássicos como os grupos sanguíneos, os antígenos dos leucócitos humanos (HLA) ou proteínas, recorrendo-se a metodologias imunológicas. Todavia, os resultados obtidos eram muito limitados e em diversas situações seriam mesmo inconclusivos (Butler, 2005; Carracedo e Sánchez-Diz, 2005; Pinheiro, 2010).

A introdução dos marcadores de ADN, assim como a rápida progressão das tecnologias associadas à sua análise contribuiu, atualmente, para a obtenção de resultados ao fim de algumas horas em amostras outrora impossíveis de estudar. As suas vantagens são inúmeras, desde a exígua quantidade de ADN necessário para se adquirir uma boa base informativa, mesmo em amostras muito degradadas. Assim, com o desenvolvimento da PCR, é possível obter um elevado número de cópias da molécula de ADN em amostras com quantidades muito ínfimas de material celular interno. Paralelamente foram introduzidos *kits* comerciais que contribuem para uma maior rapidez e eficácia na obtenção de resultados (Butler, 2005, 2012; Carracedo e Sánchez-Diz, 2005; Pinheiro, 2010).

O estudo do material genético para fins forenses é realizado com recurso a estes mesmos *kits* comerciais, com marcadores genéticos. Estes, são constituídos por reagentes e *primers* correspondentes a sequências de ADN conhecidas, com uma

determinada localização conhecida, e uma vez conhecido o padrão de transmissão à descendência, permitem a identificação genética individual.

Os marcadores utilizados na genética forense apresentam grande diversidade, sendo os de maior relevância classificados em STRs (*Short Tandem Repeats*), InDel (Inserção/Deleção) e SNPs (*Single Nucleotide Polymorphisms*). Os STRs são bastante abundantes no genoma humano, distribuídos pelos cromossomas sexuais e autossômicos. Estes marcadores são constituídos por sequências curtas, repetidas em *Tandem* e são as mais aplicadas nos laboratórios forenses devido à sua capacidade informativa. Os InDel, denominados como DIP (*deletion/insertion polymorphisms*) são marcadores de comprimento, resultando na inserção ou mutação de uma sequência específica de nucleótidos de ADN. Estes marcadores estão a despertar o interesse da comunidade forense nos últimos anos devido às suas vantagens, revelando-se um método laboratorial relativamente simples (Butler, 2010; Fondevila *et al.*, 2011; Goodwin, Linacre e Hadi, 2011; Pinheiro, 2008).

Os SNPs, são caracterizados pela ocorrência de uma “variação genética num único par de bases de um *locus* quando os dois alelos possuem uma frequência maior do que 1% numa população ao acaso” (Fig.3). Caso esta variação contenha um valor inferior a 1%, é considerada como mutação (Brookes, 1999; Butler, 2005; Carvalho, 2010; Pinheiro, 2010).

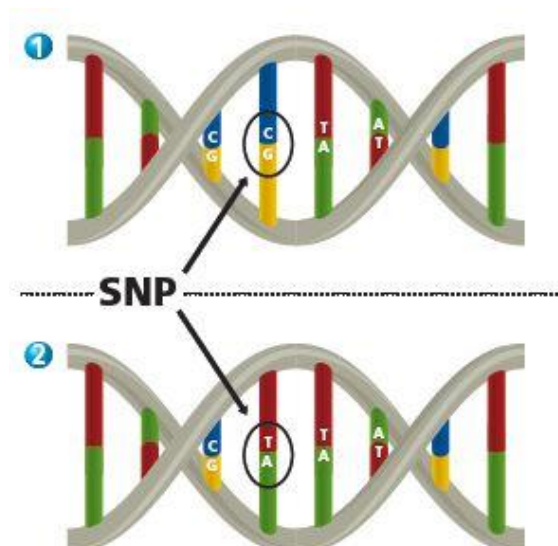


Figura 3 - Representação de um SNP. A sequência difere apenas num único par de bases. O polimorfismo é de C para T. Adaptado de Pinheiro, 2010.

Estes polimorfismos são, portanto, considerados marcadores bialélicos, caracterizados pela presença de dois alelos e só muito raramente são encontrados na forma de tri- ou tetra-alélicos pelo que não se consideram estas formas (Turner, 2008). Estão presentes com uma frequência de 1 SNP para cada 200 a 300 pb, declarando-se os marcadores com maior abundância no genoma humano (Brion *et al.*, 2004; Butler, 2012). Apesar de estarem descritos aproximadamente um milhão de SNPs, crê-se que existam cerca de 15 milhões destes marcadores genéticos. A sua frequência relativa é bastante variável, estando dependente de taxas de mutação e de recombinação associadas a cada região (Butler, 2005; Carracedo, 2005; Carvalho, 2010; Pinheiro, 2008).

1.3.1 SNPs: Aplicações Forenses

Do ponto de vista forense, os SNPs são divididos em quatro classes, de acordo com a sua aplicação: SNPs para identificação, SNPs de linhagem, SNPs de ancestralidade e SNPs de caracterização fenotípica (Butler, 2012; Pinheiro, 2010).

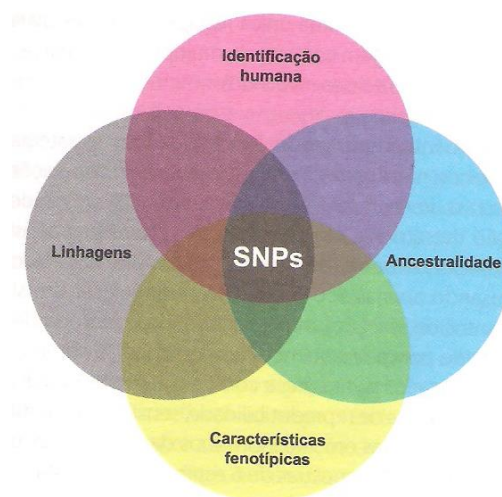


Figura 4– Esquema acerca da importância da aplicação dos SNPs na Genética Forense. Adaptado de Pinheiro, 2010.

Os SNPs para identificação humana, podem ser utilizados quando uma amostra de ADN apresenta alguma degradação e estão presentes nas regiões autossômicas do genoma. Estes marcadores são de elevada importância quando os resultados dos STRs não são conclusivos, contribuindo para complementar a informação obtida (Goodwin, Linacre e Hadi, 2011; Onori *et al.*, 2006). A estes casos de identificação podem acrescer, como já referido anteriormente, os casos de parentesco complexos (Amorim e Pereira, 2005). Nos casos de acidentes de massa ou o desaparecimento de uma pessoa, poderão distinguir-se pelo poder de identificação dos restos humanos (Brion *et al.*, 2006; Budowle e Daal, Van, 2008; Butler, 2012; Jobling, 2001; Onofri *et al.*, 2006).

Os SNPs para caracterização fenotípica são os marcadores com menor informação conhecida, provavelmente por se tratarem de marcadores com aplicação recente na área da genética forense. A análise destes SNPs permite determinar características físicas como a cor dos olhos, do cabelo ou da pigmentação da pele (Carvalho, 2010; Pinheiro, 2010; Pulker *et al.*, 2007; Sobrino e Carracedo, 2005).

Em relação aos SNPs de ancestralidade, são denominados como Marcadores Informativos de Ancestralidade (AIMs) (Budowle e Daal, Van, 2008; Carvalho, 2010; Phillips *et al.*, 2007; Pinheiro, 2010). Poderão estar localizados nos cromossomas autossômicos, no cromossoma Y, sendo estes os mais utilizados, ou corresponder a variações do ADN mitocondrial. Estes polimorfismos são bastante informativos do ponto de vista geográfico, e devido à sua baixa taxa de mutação e variabilidade de frequências alélicas entre as populações adequam-se à determinação da ancestralidade dos indivíduos (Butler, 2012; Geppert e Rower, 2012; Kosoy *et al.*, 2009; Lao *et al.*, 2006; Nievergelt *et al.*, 2013; Pinheiro, 2010; Yang *et al.*, 2005).

1.3.2 SNPs do Cromossoma Y: Aplicações Forenses

Os SNPs do cromossoma Y contribuem para a identificação de linhagens. Estes marcadores bialélicos têm particularidades que os tornam fundamentais para a determinação da origem geográfica do indivíduo (Carvalho, 2010; Pinheiro, 2010; Serdar e Sema, 2009), nomeadamente a abundância no genoma, baixa taxa de mutação e herança exclusivamente paterna, não sofrendo recombinação genética (Butler, 2012; Quintáns *et al.*, 2004; Sobrino e Carracedo, 2005; Valverde *et al.*, 2013).

A identificação de linhagens paternas, assim como o estudo das relações familiares permite a reconstrução das migrações humanas e apreender qual a

origem e expansão do homem moderno (Brion *et al.*, 2006; Carvalho, 2010; Pinheiro, 2010; Wei *et al.*, 2013). Através do estudo das linhagens poderá ser possível aferir qual o grupo étnico do contribuinte de uma determinada amostra (Lessig *et al.*, 2005; Sims, Garvey e Ballantyne, 2007). Deste modo, em situações de crime em que não existe qualquer suspeito, a investigação poderá seguir um rumo, ao reduzir potenciais suspeitos, baseando-se na área geográfica de origem do contribuinte da amostra (Jobling, Pandya e Tyler-Smith, 1997; Oven, Van, Ralf e Kayser, 2011).

Os haplogrupos determinados com o estudo dos SNPs do cromossoma Y estão distribuídos numa árvore de forma hierárquica, denominada por Árvore dos Haplogrupos Binários do Cromossoma Y, definida pelo YCC (*Y Chromosome Consortium*) em 2002, revista por (Jobling e Tyler-Smith, 2003) e por (Karafet *et al.*, 2008).

O estudo destes marcadores permite compreender que os haplogrupos, ou clades, estão associados a uma região geográfica (Daniel, Walsh e Piper, 2006; Underhill e Kivisild, 2007). É, portanto, fundamental o estudo das linhagens paternas para uma mais completa caracterização dos haplogrupos.

1.4 SNPs vs STRs

A identificação de um determinado indivíduo pode ser realizada através dos STRs, SNPs e ADN mitocondrial, todos eles com recurso a células autossómicas e/ou sexuais.

Os marcadores genéticos aplicados em toda a rotina forense são os STRs, porém, em raras situações pode recorrer-se ao estudo dos SNPs para uma mais completa informação acerca da amostra em análise (Brion *et al.*, 2005).

Diversas vantagens e desvantagens estão associadas a cada um destes diferentes marcadores. Em relação aos polimorfismos de um único nucleótido, a principal vantagem consiste na análise de ADN em amostras degradadas, pois a PCR nestes casos poderá ser realizada a partir de amostras com fragmentos de ADN com um tamanho menor de 100 pb, contrariamente aos STRs que necessitam de fragmentos com aproximadamente 300 a 400 pb (Brion, 2005; Butler, 2012; Onori *et al.*, 2006). Por outro lado, permitem obter informação associada a períodos de tempo longos na história evolutiva humana, pois têm uma taxa de mutação mais baixa, 10^{-8} , em oposição à taxa de mutação dos STRs, habitualmente 10^{-3} (Amorim e Pereira, 2005; Brumfield *et al.*, 2003; Jobling, 2001). Relativamente à análise, as reações de amplificação permitem *multiplexes* com cerca de 50 SNPs, um número de marcadores mais elevado do que sucede com os *multiplexes* dos STRs. Porém, a análise com este número grandioso é necessária para obtenção do mesmo poder discriminatório que os STRs: por cada 50 a 100 SNPs detém-se a mesma informação que 10 a 16 STRs (Butler, Coble e Vallone, 2007). É de salientar que na rotina forense, os *multiplexes* contêm 15 STRs, estando em desenvolvimento *kits* que permitem o estudo de um número superior numa única análise (Brion *et al.*, 2004; Butler, 2005, 2012; Chakraborty *et al.*, 1999; Gill, 2001; Goodwin, Linacre e Hadi, 2011; Sanchez, Børsting e Morling, 2005; Sims, Garvey e Ballantyne, 2007).

Outras vantagens dos SNPs dignas de registo consistem na inexistência de artefactos (*stutters*) durante a eletroforese capilar, o que permite simplificar a análise de perfis obtidos. Esta análise poderá ser realizada de forma mais automatizada visto a separação baseada no tamanho dos produtos amplificados não ser necessária. Sendo marcadores bialélicos, para uma validação do estudo será

necessário um menor número de amostras para calcular as frequências dos alelos nas populações. Por fim desempenham uma função relevante como marcadores de linhagem, de ancestralidade e como marcadores para estudo das características físicas de um determinado indivíduo (Carracedo, 2005; Carvalho, 2010; Pinheiro, 2010; Sanchez *et al.*, 2006; Sobrino, Brión e Carracedo, 2005).

Contudo, como estes marcadores são bialélicos, os que traz dificuldades em situações de mistura. Não há clara distinção entre um verdadeiro homozigótico, de um heterozigótico ou ainda de um *locus* com uma homozigotia e uma heterozigotia (Butler, 2005; Carvalho, 2010).

Na tabela 1, estão referidas as principais características entre STRs e SNPs.

Tabela 1– Comparação das diferentes características dos marcadores genéticos entre SNPs e STRs.

Caraterísticas	STRs	SNPs
Ocorrência no genoma humano	~ 1 em cada 15 Kb	~ 1 em cada 1 Kb
Poder informativo	Elevado	Baixo (cerca de 20 a 30% dos STRs)
Tipo marcador	Multialélico	Bialélico
Nº de alelos por marcador	Entre 5 a 20	2
Maior vantagem na genética forense	Taxa de sucesso elevada na detecção e interpretação de misturas	Taxa de sucesso elevada na análise de amostras degradadas

Em conclusão, os STRs serão sempre fundamentais para a perícia na área forense mas os SNPs têm revelado a sua importância para o estudo da expansão do homem moderno, através da classificação destes marcadores bialélicos na ancestralidade e no estudo das linhagens (Brion, 2005; Brion *et al.*, 2005; Butler, 2012).

1.5 Árvore dos Haplogrupos Binários do Cromossoma Y

A herança estritamente paterna e a ausência de recombinação são dois fatores que permitem o cromossoma Y ser um ótimo reconstituído da origem e evolução do homem moderno (Jobling, 2001).

A distribuição geográfica dos haplogrupos tem um ancestral comum, permitindo a determinação das diferentes linhagens paternas da população que foram surgindo ao longo dos séculos, através da acumulação de mutações na sequência de ADN do cromossoma (Oven, Van *et al.*, 2013). As variações que vão surgindo, a composição genética da linhagem paterna e os eventos de expansão e migração são os fatores determinantes para a definição de haplogrupos de determinada população e a variabilidade entre elas (Jobling, 2001; Schull, 1979; Underhill *et al.*, 2001).

De acordo com alguns estudos, é credível que a origem do homem moderno seja explicada pela teoria *Out Of África*, onde é admitida a hipótese de o homem ter origem em África e posteriormente ter expandido para a Europa, Ásia e Oceânia, ocorrendo uma evolução da população e o surgimento de novos haplogrupos (Myres *et al.*, 2011). Esta migração ocorreu sempre que as condições climáticas o permitiam. Esta teoria é sustentada pelo facto da existência de diversos fósseis humanos em diversos estados de evolução do homem *Homo*, em diversos lugares do planeta com datas bastante distintas. As linhagens mais antigas da árvore, A e B,

têm uma distribuição restrita a África, evidenciando que foi a partir deste continente que se sucedeu a expansão do homem moderno (Jobling e Tyler-Smith, 2003; Underhill e Kivisild, 2007).

Com a teoria *Out Of África* existiram duas épocas distintas de migração de maior importância, uma na direção da Oceânia, passando posteriormente para a Ásia (Japão) e a outra terá sido em direção à Europa e Ásia. Ambas resultaram em clades distintas da Árvore e por conseguinte definem haplogrupos diferentes (Brion *et al.*, 2004; Muro *et al.*, 2011; Semino *et al.*, 2000; Sims, Garvey e Ballantyne, 2007).

Para uma correta interpretação e uniformização dos resultados, foi criada uma nomenclatura uniforme através do YCC (*Y Chromosome Consortium*) que reuniu em 2002 e definiu um sistema hierárquico com uma terminologia, baseado nos marcadores polimórficos binários do cromossoma Y (Jobling, 2001; The Y Chromosome Consortium, 2002). Este sistema foi denominado como Árvore de Haplogrupos Binários do Cromossoma Y, publicado em 2002, revista por Jobling e Tyler-Smith (Jobling e Tyler-Smith, 2003), e mais tarde por Karafet *et al.* (Karafet *et al.*, 2008). Esta nomenclatura recente contém cerca de 311 haplogrupos, incorporando perto de 600 polimorfismos binários e encontra-se dividida em Clades (nomeadas de A a T) e em Subclades, às quais se acrescentam números.

Com a caracterização dos SNPs pertencentes ao cromossoma Y, é possível associar as migrações iniciais à distribuição atual de determinado haplogrupo. O desenvolvimento de bases de dados com dados estatísticos fiáveis e com um número mais elevado de polimorfismos binários estudados permite uma nomenclatura clara e universal, resultando numa mais fácil interpretação dos

resultados obtidos assim como a sua validação para uso de todos os laboratórios forenses (Carvalho, 2010; The Y Chromosome Consortium, 2002).

Para um determinado estudo acerca de determinada população, deverá primeiro verificar-se na árvore do YCC qual o haplogrupo mais provável para uma correta seleção de Y-SNPs. Evita-se assim, gastos desnecessários de tempo e recursos. Quantos mais marcadores forem selecionados para o *multiplex*, mais completa será a caracterização da população (Brion *et al.*, 2004).

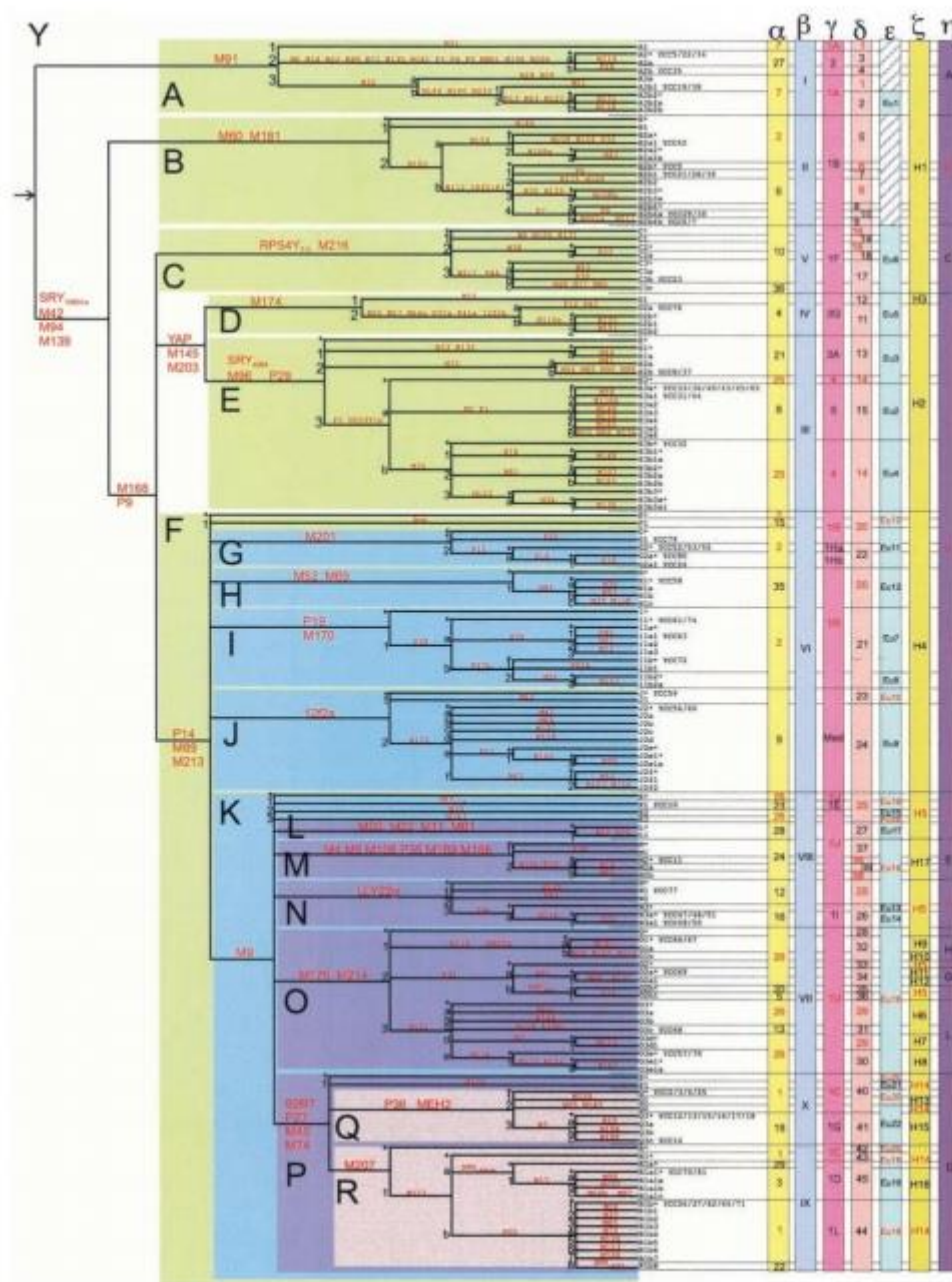


Figura 5 - Árvore dos haplogrupos binários do cromossoma Y proposta pelo *Y Chromosome Consortium* em 2002 (The Y Chromosome Consortium, 2002). Atualmente contém cerca de 311 haplogrupos, reunindo 600 polimorfismos binários. Adaptado de Y Chromosome Consortium, 2002.

1.6 Populações em estudo

1.6.1 Cabo Verde

A república de Cabo Verde é um arquipélago africano constituído por dez ilhas e oito ilhéus, de acordo com a direção do vento que atinge as mesmas. De acordo com dados dos censos de 2010, esta república é habitada por cerca de 491.875 habitantes, (Instituto Nacional de Estatística de Cabo Verde, 2010).

A região do Barvalento é formada pelas ilhas Santo Atão, São Vicente, Santa Luzia, São Nicolau, Sal e Boa Vista. A região do Sotavento é composta pelas Ilhas de Maio, Santiago, Fogo e Brava (Lopes, 2011).

O arquipélago teve origem vulcânica e apresenta 4033 km² de extensão, situando-se a cerca de 500 km do Senegal. A capital é Praia, cidade mais povoada, na ilha de Santiago (Organização Internacional para as Migrações, 2009).

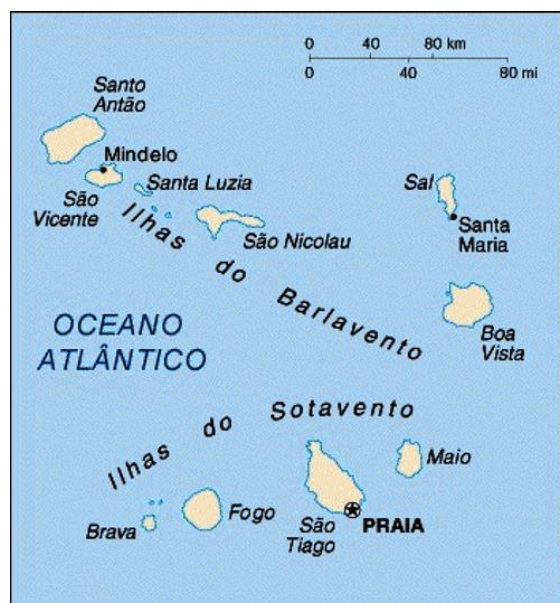


Figura 6 - Mapa de Cabo Verde, onde são visíveis as ilhas do Barlavento e as do Sotavento. Adaptado de: <http://unesdoc.unesco.org/images/0013/001330/133020por.pdf>

A sua língua oficial é o português sendo o *crioulo* cabo-verdiano o dialeto local, considerada como língua não-oficial.

Estas ilhas foram descobertas em 1460 (Fernandes *et al.*, 2003), pelo português Diogo Gomes e pelo genovês *António da Noli* em 1460, ao serviço do Infante D. Henrique que promovia a expansão do Império Colonial Português.

As primeiras ilhas identificadas foram as do Sotavento, posteriormente em 1462, por parte de Diogo Afonso, foram encontradas as restantes ilhas pertencentes ao Barlavento. Segundo os documentos da época, este arquipélago seria desabitado, permitindo a colonização por parte dos portugueses, em primeiro lugar, e mais tarde dos africanos (Brehm *et al.*, 2002).

Na ilha de Santiago situou-se a primeira povoação, designadamente em Ribeira Grande. Esta povoação era constituída por europeus exploradores que teriam vindo verificar as condições desta terra desocupada. Só no século XVIII a colonização estaria estendida a todas as ilhas do arquipélago (Brehm *et al.*, 2002).

Com a ocupação dos europeus num número muito reduzido, D. Fernando permitiu o comércio de escravos vindo, principalmente, da Guiné e do Norte de África. Assim, os primeiros habitantes de Cabo Verde foram homens pertencentes à nobreza portuguesa, judeus não convertidos, condenados ou exilados vindos de Portugal e os escravos com origem africana (Grassi, 2006). Esta combinação resultou num cruzamento entre o colono branco e a escrava negra e como consequência, a formação do nativo cabo-verdiano, denominado como *mestiço* ou *crioulo* (Brehm *et al.*, 2002).

A sua posição geoestratégica entre os três continentes permitia a prestação de serviços associada ao tráfico de escravos, que seriam posteriormente exportados para outras colônias, e a todo o comércio de mercadorias que seriam taxadas por cada navio que atracava nestas ilhas. Resultou então na expansão deste país (Grassi, 2006).

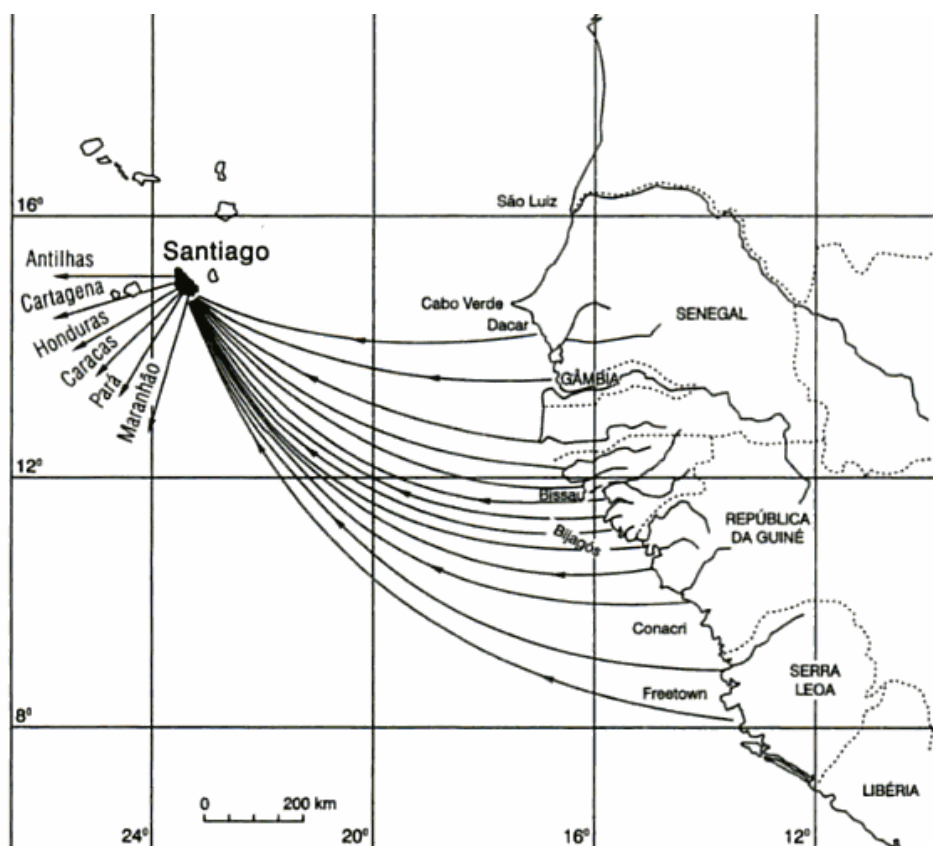


Figura 7 - Rota do tráfico de escravos. Partiam de vários pontos da costa africana com um ponto de paragem – Cabo Verde - daí seguiam para vários pontos do continente americano. Adaptado de: <https://tracingafricanroots.wordpress.com/maps/upper-guinea>.

O indivíduo cabo-verdiano pertence a um povo característico de emigração. A sua longa história migratória tem-se verificado ao longo de vários séculos, resultando numa dispersão em mais de 25 países, distribuídos pelos cinco

continentes, formando diversas comunidades (Góis, 2002; Organização Internacional para as Migrações, 2009).

A primeira dispersão ocorreu durante a escravatura, uma emigração forçada que ocorreu no século XV. Porém, o conjunto de fatores como a seca, o clima com pluviosidade irregular e terrenos inférteis levam a que o povo de Cabo Verde seja denominado como o eterno emigrante (Carling e Batalha, 2008).

Na história cabo-verdiana, fazem parte três distintas fases de grandes fluxos migratórios (Góis, 2002; Organização Internacional para as Migrações, 2009):

- 1ª Fase: Decorreu no final do século XIX e parte do século XX. Resulta com uma saída para os Estados Unidos da América assim como para o Império Colonial Português: Brasil, Angola, Moçambique e Portugal.

- 2ª Fase: Decorreu desde os anos 20 até ao fim da IIª Guerra Mundial. Nesta época ocorreu uma diminuição da ida para a Europa, sendo o maior fluxo para a América Latina (Brasil, Argentina). A ida para as colónias portuguesas continuou num elevado número.

- 3ª Fase: Decorreu desde o fim da IIª Guerra Mundial até à atualidade. O principal destino foi a Europa (Portugal, Holanda e Itália são os principais, seguidos de Espanha, Alemanha e França). Houve novamente um aumento da sua saída para os EUA.

A emigração ocorre de forma distinta conforme a ilha do arquipélago que habita. Se for natural de S. Vicente ou S. Atão irão provavelmente para o Noroeste da Europa. Se residir na ilha de Santiago, irão para Portugal. Quanto às ilhas ocidentais, Fogo e Brava, a probabilidade de ir para os EUA é maior. Estes hábitos devem-se às linhas de rede migratórias que construíram rotas das marinhas

mercantes: umas europeias e outras norte-americanas, influenciando a emigração até à atualidade (Grassi, 2006).

Atualmente a população de Cabo Verde é a segunda maior comunidade em Portugal. Segundo os censos de 2011 (Instituto Nacional de Estatística, 2012) estão a residir em Portugal cerca de 38.895 cabo-verdianos o que representa 10% da emigração total do país. De acordo com os dados do Serviço de Estrangeiros e Fronteiras, em 2013, habitavam 42.301 cabo-verdianos. Concentram-se preferencialmente na zona sul do país, estando em maior número na região de Lisboa, península de Setúbal e Algarve (Serviço de Estrangeiros e Fronteiras (SEF), 2014).

Em Portugal são o grupo mais representativo, pois para além do número oficial de emigrantes, teremos de ter em atenção os indivíduos que já se naturalizaram portugueses e os que obtiveram nacionalidade portuguesa anteriormente a 1975, por pertencerem a uma colónia portuguesa (Góis, 2002).

1.6.2 Portugal

Portugal é um país localizado a sudoeste da Europa. Situa-se na Península Ibérica conjuntamente com Espanha, com quem faz fronteira. Tem duas regiões autónomas: Madeira e Açores ao largo do Oceano Atlântico Norte.

O país tem uma área total de 92.090 km², situando-se a sua capital, Lisboa, no centro do país. A região de Lisboa tem cerca de 547.733 habitantes num total de 10.562.178 residentes em Portugal (Instituto Nacional de Estatística, 2011).



Figura 8 - Mapa de Portugal, com representação dos arquipélagos Açores e Madeira. Adaptado de: <http://www.mapadeportugal.net/indicedistritos.asp>.

A Nação europeia, que em 1139 se funda como Reino de Portugal durante a reconquista cristã, deixando para trás o condado portugalense que lhe deu origem, consegue a independência em 1143, tornando-se o país mais antigo da Europa.

Porém, muito anterior a esta época, existem vestígios de presença humana da época do Neolítico e da idade do bronze. Esta região foi também ocupada por fenícios e celtas, resultando numa mistura com os povos nativos que aqui habitavam à época (Duarte *et al.*, 1999).

No século III a.C. sucedeu-se a invasão dos romanos na península ibérica. Posteriormente, foi ocupado por povos germânicos (suevos e visigodos) e mais tarde por muçulmanos (mourões e árabes) provenientes do norte de África (Gonçalves *et al.*, 2005).

Mais tarde, no século XV, começa a Era dos Descobrimentos com a expansão do Reino de Portugal por todo o mundo: África, América e Ásia. Foi o primeiro Império Global da História e o mais duradouro, com 600 anos de existência (Mattoso, 2000).

A emigração do povo português não é uma característica adquirida recentemente. Com a descoberta e colonização de diversas terras na época dos Descobrimentos, ocorreu um grande fluxo migratório do género masculino com o objetivo de iniciar uma nova vida. Alguns eram judeus não convertidos assim como condenados ou exilados, pelo que não voltaram ao seu país de origem (Gonçalves *et al.*, 2005).

Os números de emigrantes foram-se mantendo estáveis até aos finais dos anos 60 e início dos anos 70 do século XX: ocorre um grande movimento de portugueses para a Europa, especialmente França, Suíça ou Alemanha. É de salientar que Portugal vivia numa ditadura e o povo estava a manter-se na pobreza. A maioria dos emigrantes seria homens na idade ativa de trabalho e solteiros.

A partir dos anos 90 do século passado, compreendeu-se que Portugal para além de ser um país de emigrantes, estava a tornar-se um país de imigrantes (De Oliveira, 2007; Rebelo dos Santos *et al.*, 2010). Este súbito aumento presumivelmente estará associado à entrada na CEE em 1 de Janeiro de 1986 e no Espaço de *Schengen* em 1991. A maioria dos imigrantes é de origem lusófona, africanos e brasileiros. Nesta época também ocorreu um rápido aumento da imigração proveniente da Europa de Leste (Ucrânia, Rússia, Moldávia) (Góis, 2002; Rebelo dos Santos *et al.*, 2010).

1.7 Objetivos do Estudo

Os cabo-verdianos são a segunda maior comunidade de imigrantes a residir em Portugal, concentrando-se principalmente em Lisboa e região sul de Portugal. Assim, é pertinente proceder à caracterização desta população através dos marcadores bialélicos, SNPs. Por outro lado, também é importante proceder à caracterização da população caucasiana portuguesa, de modo a que a informação obtida possa ser utilizada na área forense e criminal de identificação genética, passando a ser exercida na rotina forense do serviço de biologia e genética forense na delegação sul do INMLCF, IP.

Assim, os principais objetivos deste trabalho consistem em:

- Estudar indivíduos da população de Cabo Verde e de Portugal com Y-SNPs para caracterização dos seus haplogrupos;
- Realizar um estudo populacional da ancestralidade das duas populações;
- Calcular distâncias genéticas entre os dois grupos populacionais e verificar se existem diferenças significativas entre estes dois grupos populacionais.

MATERIAL E MÉTODOS

2.1 Colheita e Armazenamento de Amostra

Neste estudo, foram selecionados 150 indivíduos do género masculino, saudáveis e não relacionados entre si. Desta amostra, 75 são de origem cabo-verdiana e outros 75 de raça caucasiana, de origem portuguesa. Em cada uma das nacionalidades foi verificada a ancestralidade, ou seja, que os pais destes intervenientes também tinham a mesma nacionalidade, de modo a garantir a veracidade deste estudo. Estes indivíduos são residentes no Sul de Portugal, e dirigiram-se ao Serviço de Biologia e Genética Forense da Delegação Sul do Instituto de Medicina Legal e Ciências Forenses, IP. para a respetiva colheita, no âmbito de processos de investigação de parentesco.

A colheita da amostra, após consentimento informado, foi efetuada por punção capilar em cartão Whatmann® devidamente identificado. Estas amostras biológicas são secas à temperatura ambiente e depois armazenada num envelope devidamente identificado prevenindo a ocorrência de qualquer tipo de contaminação.

Durante todo o processo, desde a colheita até ao momento do estudo da amostra, todas as normas de boas práticas laboratoriais devem ser mantidas de modo a garantir a integridade biológica da mancha de sangue (Butler, 2005; Petkovski *et al.*, 2006).

2.2 Extração de ADN

A análise de uma amostra biológica com o objetivo de estudar o seu ADN pressupõe a extração e purificação deste.

O material genético que se pretende estudar encontra-se no interior da célula, mais concretamente no seu núcleo, estando protegido por proteínas para que não ocorra qualquer tipo de modificação devido ao ambiente celular envolvente (Butler, 2005). Assim, para proceder à sua análise, é necessário separar o ADN de todas as proteínas envolvidas bem como do restante material celular, de modo a que não existam elementos contaminantes que possam interferir com a respetiva análise.

A extração de ADN é realizada em três fases: 1. Lise celular que consiste na destruição dos componentes subcelulares; 2. Separação das moléculas de ADN do restante material; 3. Purificação de material genético para a fase final do seu estudo (Butler, 2010).

A separação e purificação do material genético são fases muito importantes pois durante os procedimentos seguintes, como amplificação, podem surgir problemas devido à presença de proteínas ou qualquer outro material inibidor. Estes produtos podem interferir no emparelhamento do ADN- Molde com o *primer*, realçando a importância de uma correta extração, a qual poderá ser feita sob diferentes métodos (Butler, 2012).

Neste trabalho foram estudadas amostras cujo ADN foi extraído pelo método de Chelex®100. Este método, introduzido na comunidade forense em 1991, baseia-se na utilização de uma resina quelante com o objetivo de retirar o ADN do interior do núcleo da célula (Bio-Rad Laboratories, 2000; Walsh, Metzger e Higuchi, 1991).

O Chelex®100 é, portanto, uma resina quelante (Sims, Garvey e Ballantyne, 2007), composta por copolímeros de estirenodivinil benzeno que contém iões iminodiacetato emparelhados e que vão funcionar como grupo quelante de iões metálicos como o caso do magnésio. Como consequência da remoção do magnésio as nucleases, enzimas responsáveis pela rutura de ligações fosfodiéster e respetiva degradação do ADN, são inativas e as moléculas de ADN permanecem protegidas (Brisighelli *et al.*, 2006).

Esta resina é adicionada diretamente à amostra, conseguindo atrair os iões magnésio, funcionando como um íman, impedindo a degradação do ADN. A conjugação de Chelex 5% com a lavagem prévia das amostras e a exposição a altas temperaturas terá como resposta a destruição das membranas celulares e a desnaturação do ADN, que resulta em moléculas de cadeias simples em vez da dupla cadeia, podendo ser amplificada por PCR (Walsh, Metzger e Higuchi, 1991).

Este método de extração é bastante utilizado nos laboratórios forenses pois a sua capacidade de remoção dos inibidores de PCR e do grupo heme das proteínas, o facto de ser adicionada diretamente à amostra e de não haver troca de tubos reduzem bastante o potencial de contaminação, revelando-se um método de eleição (Butler, 2005, 2010, 2012).

No Anexo I, está presente o protocolo de extração de ADN a partir de manchas de sangue por Chelex®100.

2.3 Quantificação de ADN

Na área forense são estudadas amostras problema que estão sujeitas a fatores exógenos que contribuem não só para a degradação do ADN como também para a sua contaminação, um problema muito comum nos laboratórios forenses. Por outro lado, mesmo em amostras de referência, o conhecimento prévio da quantidade de ADN presente na amostra é importante em métodos muito sensíveis e morosos, como por exemplo nas práticas laboratoriais utilizadas no âmbito deste trabalho (Butler, 2010, 2012).

Assim, antes de se prosseguir para as restantes etapas, a quantificação é crucial pois permite verificar se existe quantidade mínima de ADN na amostra extraída para uma subsequente amplificação, principalmente em amostras degradadas ou com um baixo número de cópias desta molécula (Alonso *et al.*, 2003; Goodwin, Linacre e Hadi, 2011; Heid *et al.*, 1996). Este método também permite obter informação sobre a presença de inibidores, nomeadamente o grupo heme da hemoglobina, (presente nas amostras de sangue) e que quando extraídos em conjunto com o ADN vão interferir com a amplificação dos marcadores pretendidos (Andréasson, Gyllensten e Allen, 2002; Nicklas e Buel, 2003).

A PCR em Tempo-Real (*RT-PCR*), método descrito em 1993, é um procedimento que permite quantificar o ADN de uma determinada amostra através da amplificação em tempo real (Higuchi *et al.*, 1993). A RT-PCR é realizada com recurso a *kits* desenvolvidos para esse efeito e ainda com um equipamento apropriado designado por *ABI PRISM 7500 Real - Time PCR Systems (Applied Biosystems)*.

No âmbito da genética forense, foram desenvolvidos *kits* de quantificação. Neste estudo, foi utilizado para a quantificação das amostras o *kit Quantifiler® Human DNA Quantification Kit (Applied Biosystems)* que permite quantificar o ADN nuclear humano presente, o qual foi amplificado e detetado por PCR em tempo real no equipamento *ABI Prism 7500*, com a utilização do software *HID Real Time PCR Analysis Software versão 1.2*.

Este *kit* permite a quantificação através da amplificação do locus do intrão do gene da transcriptase reversa da telomerase do cromossoma 5. A nível qualitativo também deteta a presença de inibidores de PCR com a introdução de um controlo positivo interno – IPC, que tem como função verificar a capacidade de amplificação das amostras que estão a ser analisadas (Applied Biosystems, 2012).

O protocolo para a quantificação nuclear de ADN com o *Quantifiler® Human Quantification Kit* encontra-se no Anexo II.

2.4 Genotipagem de Y- SNPs

Nos últimos anos têm sido desenvolvidas várias tecnologias para o estudo adequado destes marcadores, porém a comunidade forense ainda não chegou a um consenso sobre qual a mais eficiente, tendo, porém, definido que deverá ter uma elevada sensibilidade, boa relação custo-eficiência, automatização e boa capacidade de resposta-tempo (Bríón *et al.*, 2006; Jobling e Tyler-Smith, 1995; Jobling, 2001; Nilsson *et al.*, 2006; Pati *et al.*, 2004).

O estudo dos Y-SNPs desta dissertação foi efetuado por minisequenciação com o *kit SNaPShot™ Multiplex Kit (Applied Biosystems)*. É um método que se tem

revelado mais adequado e eficiente para a análise de SNPs, neste caso em particular, referentes ao cromossoma Y (Pati *et al.*, 2004; Schell *et al.*, 2006).

A genotipagem de Y-SNPs por este método caracteriza-se por três fases, definidas na tabela 2: Amplificação dos SNPs do Cromossoma Y, a extensão dos *primers* e por fim a sua detecção e análise (Brion *et al.*, 2004). Entre a fase da amplificação e da extensão, é realizada uma etapa de purificação com o conjunto de enzimas *ExoSAP-IT*[®] (*Usb*[®]) que remove não só os *primers* não incorporados, como também os *dNTPs* que se encontram livres (Brion *et al.*, 2004; Brookes, 1999; Gusmão *et al.*, 1999).

Tabela 2– Resumo dos métodos laboratoriais para a análise dos SNPs do Cromossoma Y.

Método Laboratorial	
Amplificação de Y-SNPs	◦ Reação de PCR da amostra de ADN com os <i>Primers</i> presentes no Multiplex.
Extensão de <i>Primers</i>	◦ Reação de minisequenciação da amostra já amplificada e purificada com o <i>Kit SNaPShot</i> [™] .
Deteção e Análise	◦ Aplicação da amostra num no sequenciador <i>ABI Prism</i> [®] 3130 <i>Genetic Analyzer</i> para eletroferese capilar e análise com o <i>Gene Mapper</i> [®] <i>ID Software</i> versão 3.2.1.

Posteriormente, após a extensão e antes da detecção, os *ddNTPS* não incorporados são inativos pela enzima *SAP*[®] (*Roche*[®]) (Brion *et al.*, 2004).

Os SNPs selecionados para esta investigação, e de acordo com *Brión et al.* (Brion *et al.*, 2004), foram os seguintes: SRY1532, M70, M173, M22, P25, 92R7, Tat, M213 e M9, referenciados na tabela 3, estudados em conjunto numa reação *Multiplex*. De acordo com estudos já devidamente publicados e em concordância com a sua localização na Árvore de Haplogrupos Binários do Cromossoma Y, estes 9 SNPs são os mais oportunos para uma correta identificação dos haplogrupos destas duas populações (Brion *et al.*, 2004; Cruciani *et al.*, 2010; Sims, Garvey e Ballantyne, 2007; The Y Chromosome Consortium, 2002).

Na tabela 3, além da designação do SNP em estudo, está também indicada qual a base nucleotídica detetada e a mutação correspondente, assim como o seu respetivo tamanho.

Tabela 3 - Seleção de Y-SNPs distribuídos no multiplex, com informação da mutação e respetiva deteção, segundo *Brión et al.* (2004).

Y-SNP	Tamanho	Deteção	Mutação
M22	20	A – G	A → G
P25	26	C – AC	C → A
92R7	28	TC – T	G → A
SRY1532	30	A – G	A → G
M70	34	T – G	A → C
M173	34	A – C	A → C
Tat	42	A – G	T → C
M213	45	A – G	T → C
M9	48	C – G	C → G

2.4.1 Amplificação de ADN

A técnica que foi utilizada para esta investigação teve a sua descoberta nos anos 80, e é denominada por PCR – Reação de Polimerização em Cadeia (Mullis *et al.*, 1992). Foi considerada uma tecnologia revolucionária pois tem a capacidade de multiplicar, em pouco tempo, milhares de cópias de um fragmento de ADN específico através de uma sequência-molde (Butler, 2005, 2010).

Uma característica essencial para este êxito é o facto da amostra de ADN poder estar degradada e consequentemente terá um número muito ínfimo de cópias de ADN. Com a técnica de PCR, esta poderá ser analisada, pois a partir de um fragmento obtém-se um número bastante elevado de cópias idênticas (Butler, 2012; Goodwin, Linacre e Hadi, 2011).

A aplicação deste método laboratorial da técnica de PCR pressupõe a utilização de *primers*, sequências iniciadoras e finalizadoras, *dNTPs*, *Taq Polimerase* (DNA Polimerase estável a temperaturas superiores a 72°C) assim como uma solução tampão e um equipamento apropriado, designado por termociclador (Carracedo e Sánchez-Diz, 2005).

Os *primers* são sequências de ácidos nucleicos complementares à cadeia original que se ligam na extremidade 5' do segmento alvo a amplificar (Fig. 9) Estas sequências vão hibridizar com a sequência do ADN – Molde no início da região que se pretende estudar, seguindo-se a adição de nucleótidos para que a cadeia molde seja copiada pela *Taq Polimerase* (Videira, 2001). A reação ocorre em ciclos com diferentes temperaturas com um programa de amplificação apropriado a cada método (Fig.10).

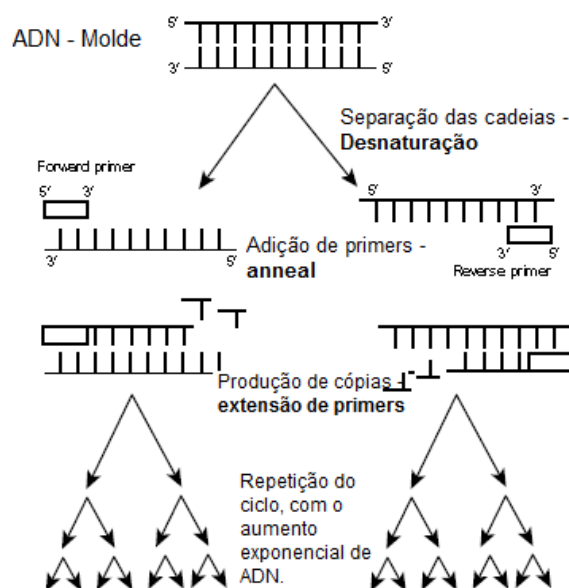


Figura 9 - Representação esquemática da reação de PCR. Adaptado de Butler, 2005.

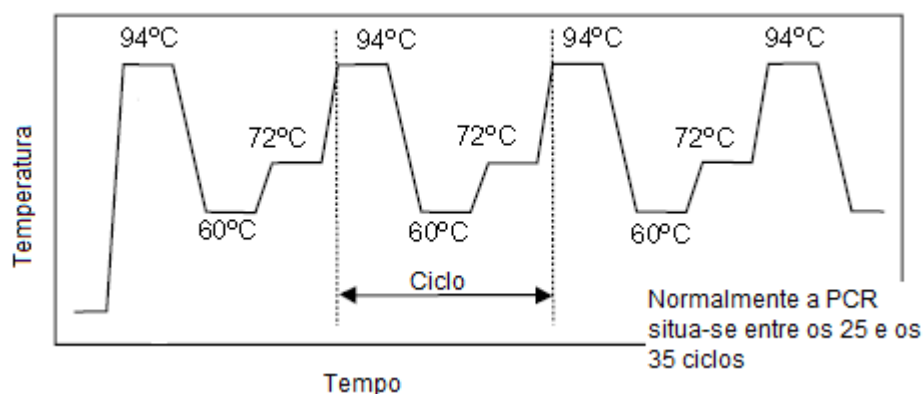


Figura 10 - Representação de ciclos de PCR. É dividido em 3 diferentes temperaturas que são repetidas entre 25 a 35 ciclos. A 94°C o ADN é desnaturado, como efeito há a separação da cadeia dupla em duas cadeias simples, em que cada uma das cadeias irá servir de molde para cópia. A 60°C sucede-se o emparelhamento dos primers com a cadeia molde de ADN. Aos 72°C a *DNA Polimerase* permite a extensão dos primers e a formação de uma nova cadeia. Adaptado de Pinheiro, 2008.

Para uma amplificação de mais do que uma região da molécula de ADN em simultâneo foi criado um *multiplex*, que se caracteriza como um conjunto de diferentes *primers* para uma polimerização em cadeia de diferentes regiões da mesma molécula (Brisighelli *et al.*, 2006).

O *kit* de amplificação de PCR utilizado para este estudo foi o QIAGEN® *Multiplex PCR Kit*, cuja composição está descrita na tabela 4, desenvolvido pela Qiagen® com o intuito de simplificar esta técnica, tornando-a também mais rápida (QIAGEN®, 2010).

Tabela 4 - Composição do *kit* comercial QIAGEN® *Multiplex PCR Kit*.

Qiagen® Multiplex PCR Kit	
2x Qiagen Multiplex PCR Master Mix	<ul style="list-style-type: none"> ◦ HotStarTaq® ADN Polimerase; ◦ Tampão Multiplex PCR; ◦ Mistura de dNTP.
Q- Solution	
Água sem RNase	.

A esta mistura são adicionados os *primers* pretendidos, SRY1532, M70, M173, M22, P25, 92R7, Tat, M213 e M9, conforme descrito por Brión *et al.* representados na tabela 5, com a concentração de 0,2 µM (Brion *et al.*, 2004).

Tabela 5 - Caracterização dos *primers* utilizados para amplificação dos Y-SNPs.

Y-SNP	Primers Diretos (5' → 3')	Primers Reversos (5' → 3')	Tamanho (pb)
M22	GCTGATAGTCCTGGTTTCCTA	TGAGCATGCCTACAGCAGAC	106
P25	GGACCATCACCTGGGTAAAGT	AGTGCTTGTCCAAGGCAGTA	121
92R7	TGCATGAACACAAAGACGTA	GCATTGTTAAATATGACCAGC	55
SRY1532	TCCTTAGCAACCATTAATCTGG	AAATAGCAAAAACTGACACAAGGC	167
M70	TCATAGCCCACTATACTTTGGAC	CTGAGGCTGGACTATAGGG	81
M173	GCACAGTACTCACTTTAGGTTTGC	GCAGTTTCCCAGATCCTGA	172
Tat	GACTCTGAGTGTAGACTTGTGA	GAAGGTGCCGTAAAAGTGTGAA	112
M213	GGCCATATAAAAACGCAGCA	TGAATGGCAAATTGATTCCA	208
M9	GCAGCATATAAACTTTTCAGG	AAAACCTAACTTTGCTCAAGC	340

O termociclador utilizado foi o *Veriti® 96-Well Thermal Cycler (Applied Biosystems)*, cujo programa de amplificação está descrito na tabela 6. O protocolo respeitante à amplificação de ADN com o *QIAGEN® Multiplex PCR Kit* está descrito no anexo III.

Tabela 6 - Programa para amplificação das amostras.

Temperatura	Duração		
95°C	15 Minutos	Incubação	
94°C	30 Segundos	Desnaturação	35 Ciclos
60°C	90 Segundos	Emparelhamento	
72°C	60 Segundos	Extensão	
72°C	10 Minutos	Extensão Final	
4°C	+ ∞		

2.4.2 Purificação com *ExoSAP-IT*[®]

Todas as técnicas aqui descritas são extremamente sensíveis e qualquer contaminação ou algum passo incorreto pode comprometer toda a investigação, podendo interferir com os resultados finais, como tal, será necessário recorrer a processos de purificação.

Depois da reação de polimerização em cadeia, nos tubos de reação estão milhares de cópias que foram produzidas, bem como *dNTPs* e *primers* que não foram incorporados assim como componentes do tampão do *kit* de amplificação (Lessig *et al.*, 2005).

Antes de se proceder à minisequenciação, o produto amplificado é purificado através de um tratamento enzimático com a adição de duas enzimas: *Exonuclease I* (*Exo I*) e a *Shrimp Alkaline Phosphatase* (*SAP*) que são comercializadas em conjunto: *ExoSAP-IT*[®] (*Usb*[®]) (Bell, 2008; Brión *et al.*, 2005). A função da *Exo I*

consiste em degradar os *primers* da cadeia e o ADN que não pertence a esta. A enzima *SAP* tem como função desfosforilar os *dNTPs* que não foram incorporados e que permanecem em solução (Fig. 11 e 12) (Bell, 2008).

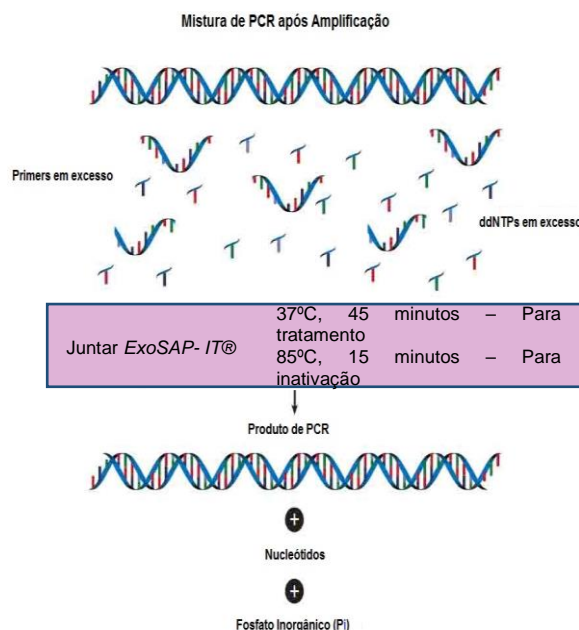


Figura 11 - Esquema de purificação com ExoSAP-IT® (usb®). Adaptado de: tudogeral.sib.uc.pt.

Estas enzimas são adicionadas diretamente à amostra tornando-se uma técnica bastante simples e vantajosa, não havendo perigo de contaminação ou perda de amostra. O método é rápido e pode ser utilizado em amostras múltiplas tornando-se bastante interessante e amplamente utilizado (Butler, 2012; Werle *et al.*, 1994).

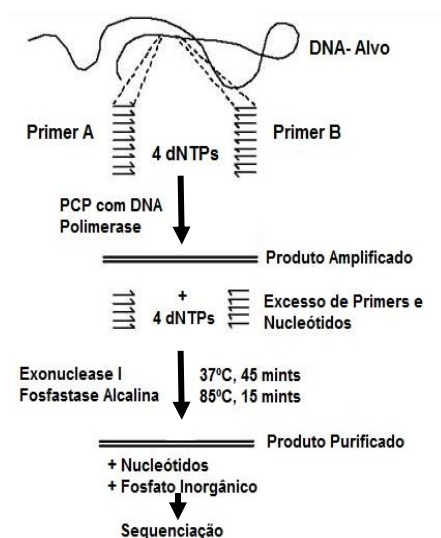


Figura 12 - Esquema geral de funcionamento da ExoSAP-IT®. Adaptado de: <http://labs.icb.ufmg.br/lbem/protocolos/exosap.html>

Este método consiste na adição de 1 µL de *ExoSAP-IT*® às amostras em estudo, incubação a 37°C durante 45 minutos enquanto ocorre o processo de purificação e de seguida inativação das enzimas a 85°C durante 15 minutos.

Este procedimento foi realizado no termociclador foi o *Veriti*®96-Well Thermal Cycler (*Applied Biosystems*). Depois deste processo, as amostras estão prontas para a técnica de minisequenciação.

2.4.2 Minisequenciação com SNaPShot™ Multiplex Kit

O kit comercial *SNaPShot*™ Multiplex Kit tem-se tornado a tecnologia mais utilizada e que mais consenso gera nos laboratórios forenses para o estudo de SNPs do Cromossoma Y (Jobling, 2001; Lessig *et al.*, 2004; Sanchez *et al.*, 2003; Schell *et al.*, 2006).

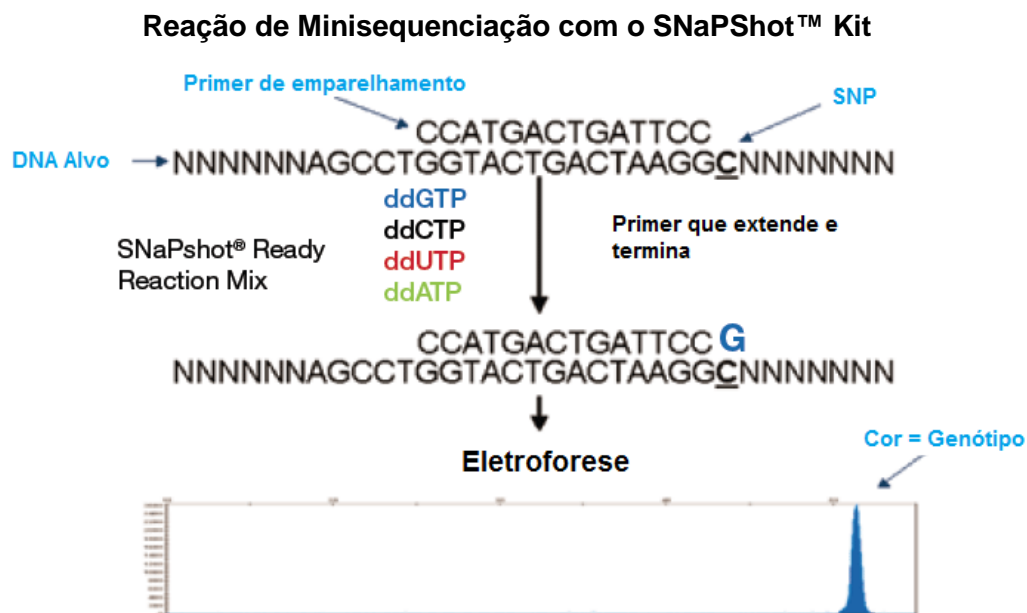


Figura 13 - Esquema geral da reação de minisequenciação com o *SNaPShot™ Kit*.

Adaptado de Applied Biosystems, 2012.

É um método de Extensão de uma Base Única (*SBE, Single Base Extension*), que analisa polimorfismos de um nucleótido – SNP. Esta técnica começa com uma minisequenciação que vai permitir a genotipagem de Y-SNPs, através da determinação da identidade da base nucleotídica que está numa determinada posição ao longo do ácido nucleico e da qual segue a sua detecção eletroforética por fluorescência (Fig. 13) (Applied Biosystems, 2010, 2012; Geppert e Rower, 2012; Sobrino e Carracedo, 2005).

A reação baseia-se numa extensão dideoxi de uma base complementar através de um *primer* não marcado. Vai ocorrer um emparelhamento ao ADN – Alvo que está adjacente ao SNP. Esta ligação acontece na presença da *DNA Polimerase* assim como *ddNTPs* marcados com fluorescência. Esta polimerase vai permitir a extensão de um nucleótido através do *primer* não marcado, que quando posicionado na sua localização correta (que será imediatamente após ao SNP) é estendido em uma única base pela adição de um *ddNTP* (Applied Biosystems, 2010; Sanchez *et al.*, 2003; Sobrino e Carracedo, 2005).

Cada *ddNTP* que está marcado com fluorescência tem um diferente fluorocromo (Tabela 7) que vai emitir um comprimento de onda distinto, o que resulta em cada base associada a uma cor e que são diferentes quando analisadas por eletroforese capilar (Mengel-Jørgensen *et al.*, 2004).

Tabela 7 - Representação dos fluorocromos e o seu respetivo ddNTP assim como a cor que é visível na eletroforese. Em relação ao ddCTP, para uma melhor observação, a cor que surge no eletroforegrama é preto no lugar do amarelo.

ddNTP	Fluorocromo	Cor
A	dR6G™	Verde
C	dTAMRA™	Amarelo (Preto)
G	dR110™	Azul
T	dROX™	Vermelho

A certificação dos diferentes tamanhos dos *primers* é efetuada através da adição de uma cauda de sequências não emparelháveis (dA, dC, dT e dG) à extremidade 5' do *primer*. Deste modo, o tamanho será diferente e não afeta a eletroforese capilar, pois o desenho dos *primers* assim como a purificação do ADN em estudo pode afetar significativamente a genotipagem (Sanchez, Børsting e Morling, 2005).

Os *primers* utilizados para a reação de minisequenciação encontram-se descritos na tabela 8, assim como a respetiva concentração usada na reação, conforme descrito por *Brion et al.* (Brion *et al.*, 2004).

Tabela 8 - Caracterização dos *primers* para a reação de minisequenciação dos Y-SNPs.

Y-SNP	<i>Primers</i> para Minisequenciação	Tam. (pb)	Conc. (μ M)
M22 (Dir)	CCGCCATTCTGGTGGCTCT	20	0.10
P25 (Dir)	CCCCCCTCTGCCTGAAACCTGCCTG	26	0.15
92R7 (Rev)	CCCCGCATGATGAACACAAAAGACGTAGAAG	28	0.20
SRY1532 (Dir)	CCCCCCTTGATCTGACTTTTTCACACAGT	30	0.15
M70 (Rev)	CCCCCCCCTAGGGATTCTGTTGTGGTAGTCTTAG	34	0.15
M173 (Dir)	CCCCCCCCCCTTACAATTCAAGGGCATTTAGAAC	34	0.20
Tat(Rev)	CCCCCCCCCCCCCCCCCCTCTGAAATATTAAATTAACAAC	42	0.20
M213 (Rev)	CCCCCCCCCCCCCCCCCCTCAGAACTTAAACATCTCGTTAC	45	0.36
M9 (Dir)	CCCCCCCCCCCCCCCCCCCCCGAAACGGCCTAAGATGGTTGAA T	48	0.25

O kit comercial *SNaPShot™ Multiplex Kit* é composto por *SNaPShot Multiplex Ready Reaction Mix* e por dois controles: *SNaPShot Multiplex Control Primer Mix* e *SNaPShot Multiplex Control Template*, representados na tabela 9. (Applied Biosystems, 2010).

Tabela 9 - Composição do kit comercial *SNaPShot™ Multiplex Kit*.

<i>SNaPshot™ Multiplex Kit</i>	
<i>SNaPshot Multiplex Ready Reaction Mix</i>	<ul style="list-style-type: none"> ◦ AmpliTaq® DNA Polimerase; ◦ ddNTPs marcados com fluorescência; ◦ Tampão de reação.
<i>SNaPshot Multiplex Control Primer Mix</i>	
<i>SNaPshot Multiplex Control Template</i>	

O termociclador utilizado para esta reação foi o *Veriti® 96-Well Thermal Cycler* (Applied Biosystems) com o programa descrito na tabela 10.

As amostras foram manuseadas segundo o protocolo utilizado para minisequenciação com o *SNaPShot™ Multiplex Kit* que se encontra descrito mais detalhadamente no Anexo IV.

Tabela 10 - Programa de PCR para a Minisequenciação.

Temperatura	Duração	
96°C 50°C 60°C	10 Segundos 5 Segundos 30 Segundos	25 Ciclos
4°C	+ ∞	

2.4.3 Purificação com SAP

Após a fase de minisequenciação, é necessária uma nova purificação de modo a que os ddNTPs que não foram incorporados durante a reação anterior sejam removidos. Deste modo, não há interferência de nucleótidos que não tenham interesse, evitando assim uma genotipagem errada (Applied Biosystems, 2010; Jobling, 2001; Mengel-Jørgensen *et al.*, 2004).

Este tratamento enzimático é feito através da adição de SAP® (Roche®) diretamente na amostra, catalisando a remoção do grupo 5' fosforil dos *ddNTPs* (Bell, 2008). É um processo bastante eficiente, rápido, de baixo custo e com uma probabilidade reduzida de contaminação visto que é adicionado 2 µL de SAP® (Roche®) diretamente no tubo da amostra posteriormente à minisequenciação (Butler, 2012; Werle *et al.*, 1994). Este procedimento (tabela 11) foi realizado no termociclador *Veriti® 96-Well Thermal Cycler* (Applied Biosystems) durante 60 minutos a 37°C para tratamento enzimático. Após esse período, as amostras são colocadas durante 15 minutos a uma temperatura de 85°C para completa inativação da enzima.

A amostra está preparada para a análise por eletroforese capilar (Sobrinho, Brión e Carracedo, 2005).

Tabela 11 - Programa descrito no termociclador para tratamento enzimático com a SAP® (Roche®).

Temperatura	Duração
37°C	60 Minutos
85°C	15 Minutos

2.5 Detecção e Análise do Produto Minisequenciado

As amostras em estudo foram sequenciadas no sequenciador automático *ABI Prism® 3130 Genetic Analyzer* de 4 capilares (Applied Biosystems, 2009, 2010)., equipamento de elevada sensibilidade, rápida injeção da amostra e análise de resultados em formato digital, com todas as vantagens que advém destas características (Applied Biosystems, 2010).

O processo de electroforese ocorre através da migração da amostra, que irá ser detetada pelo laser devido à fluorescência da base complementar à zona polimórfica do Y-SNP em estudo. Conforme mencionado anteriormente, cada ddNTP emite um comprimento de onda diferente correspondendo a uma cor. A base presente irá ser determinada pelo seu tamanho, pela localização do seu pico e pela sua cor. A cor emitida é gerada na sequenciação (Pati *et al.*, 2004).

Para que a análise das amostras seja correta recorre-se a um padrão interno de tamanho conhecido, *Internal Size Standard*, que é adicionado a cada amostra,

emitindo uma fluorescência de cor diferente dos ddNTPs, sendo laranja. Para esta investigação, o padrão foi o *Gene Scan® 120 LIZ size standard* (Applied Biosystems, 2009, 2010, 2012).

As análises foram efetuadas pelos softwares *ABI Prism® 3130xlGeneticAnalyzer Data Collection Software versão 3.0.* e *ABI Gene Mapper® ID Software versão 3.2.1* e No primeiro são definidos os filtros a que vão corresponder os diferentes comprimentos de onda. Após essa detecção de fluorescência, o software *ABI Gene Mapper® ID Software versão 3.2.1* permite a interpretação dos resultados através da intensidade dessa fluorescência, o que corresponde a uma quantidade de fragmentos de ADN (Applied Biosystems, 2009).

Para a realização da eletroforese com o primeiro software referido foi introduzido um tempo de injeção na ordem dos 15 segundos e com uma voltagem de 1.0 kVolts. O polímero aplicado foi o POP-6™, especificamente formulado para separar fragmentos de ADN com tamanho conhecido, com o tempo de execução e realização desejados.

Amplificação por PCR

Reação de minisequenciação

Eletroforese

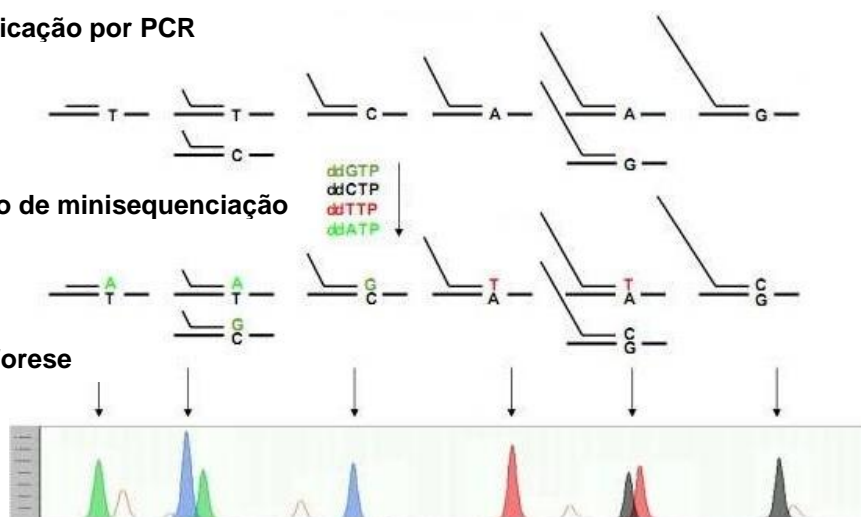


Figura 14 - Representação do funcionamento da Reação de Minisequenciação com o *SNaPSHOT™ Multiplex Kit*. Adaptado de Applied Biosystems, 2012.

A análise no sequenciador automático é realizada após uma prévia preparação das amostras, como está descrito no anexo V.

2.5.1 Detecção do Produto Sequenciado

Através do sequenciador *ABI Prism® 3130 Genetic Analyzer*, as amostras resultantes da minisequenciação marcadas com fluorescência são detetadas e separadas por eletroforese capilar.

O resultado é apresentado na forma de eletroforegrama com informação sob forma de fragmentos com diferente coloração identificando a base complementar ao local polimórfico em estudo (Applied Biosystems, 2012).

Cada um dos quatro nucleótidos está marcado com uma diferente cor: G – Azul, C – Amarelo (preto), A – Verde e T – Vermelho. Portanto, se o pico que estiver registado for de cor vermelha, sabemos que terá sido a Timina (T) a se complementar ao SNP que está a ser estudado (Fig.15) (Applied Biosystems, 2009, 2010).

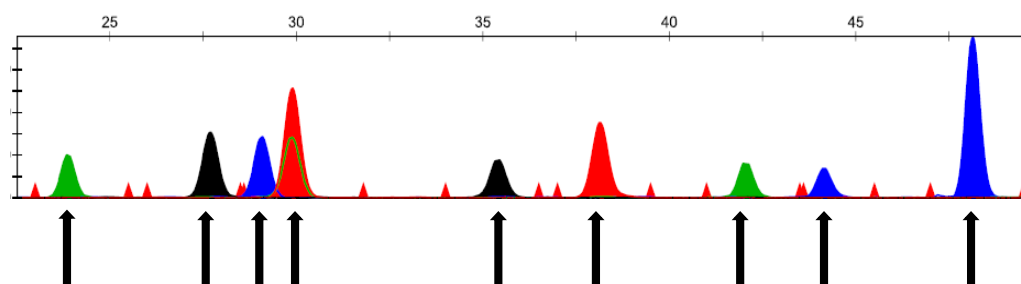


Figura 15 - Representação de um eletroforegrama com os 9 Y-SNPs em estudo.

Os picos, que representam o tamanho do fragmento de ADN, são separados através do *ABI Prism® 3130xl Genetic Analyzer Data Collection Software* versão 3.0 com a ajuda de um padrão interno de tamanho conhecido, o *Gene Scan® 120 LIZ size standard*. Este padrão é marcado com um quinto fluorocromo de cor laranja e foi desenhado especificamente para ser utilizado com o *SNaPShot™ Multiplex Kit*, tendo como função atribuir tamanho a fragmentos que se situem entre os 15 e os 120 pares de bases (Applied Biosystems, 2012).

Contém cerca de nove fragmentos de ADN de cadeia simples com tamanho de 15, 20, 25, 35, 50, 62, 80, 110 e 120 nucleótidos (Fig.16).

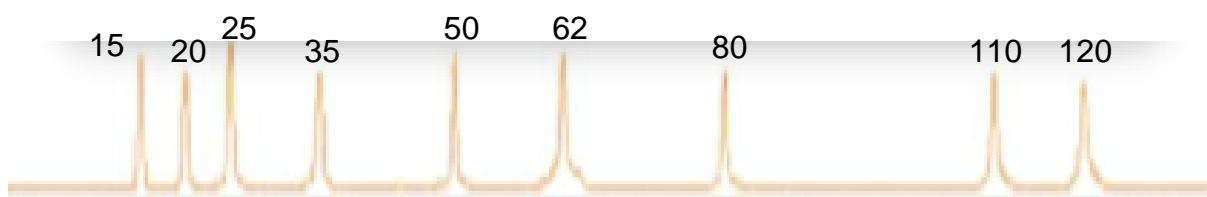


Figura 16 - Representação dos fragmentos do padrão interno *Gene Scan® 120 LIZ size standard*.

2.6 Análise estatística, designação de Haplogrupos e Árvore Filogenética

Por fim, depois de analisado e identificado cada pico referente a cada Y-SNP, prossegue-se para a atribuição do haplogrupos de cada amostra em estudo. O haplogrupo foi determinado segundo a Árvore de Haplogrupos Binários do Cromossoma Y, definido pelo YCC (The Y Chromosome Consortium, 2002), revista

por Jobling e Tyler-Smith (Jobling e Tyler-Smith, 2003) e atualizada por Karafet *et al.* (Karafet *et al.*, 2008)

O estudo das frequências dos haplogrupos foi feito com o *software IBM – SPSS versão 21.0*, programa de organização de dados e de análise estatística.

A distância genética foi calculada através do *Arlequin versão 3.5.2*, *software* que trabalha com testes estatísticos a partir das frequências calculadas e que identifica características genéticas e demográficas acerca da população em estudo.

RESULTADOS E DISCUSSÃO

Depois da etapa de extração de ADN, amplificação, minisequenciação com *SNaPShot™ Multiplex Kit* e detecção por eletroforese capilar no *ABI Prism® 3130 Genetic Analyzer* concluída, os resultados obtidos foram analisados, determinando-se qual a base nucleotídica presente em cada amostra.

Foram analisados os polimorfismos de 9 Y-SNPs de 150 indivíduos: 75 de origem caucasiana, especificamente portuguesa e 75 cabo-verdianos.

Deste modo, foi possível determinar qual a variação nucleotídica e por comparação com a Árvore de haplogrupos binários do Cromossoma Y, proposta pelo YCC (The Y Chromosome Consortium, 2002), revista por *Jobling e Tyler-Smith* (Jobling e Tyler-Smith, 2003) e por *Karafet et al.* (Karafet et al., 2008), concluiu-se qual o haplogrupo de cada indivíduo das duas populações em estudo. A caracterização mais detalhada do haplogrupo de cada indivíduo encontra-se no anexo VI, tabelas 14 e 15.

Por fim, prosseguiu-se ao cálculo das frequências de cada haplogrupo, assim como à análise das distâncias genéticas entre as duas populações.

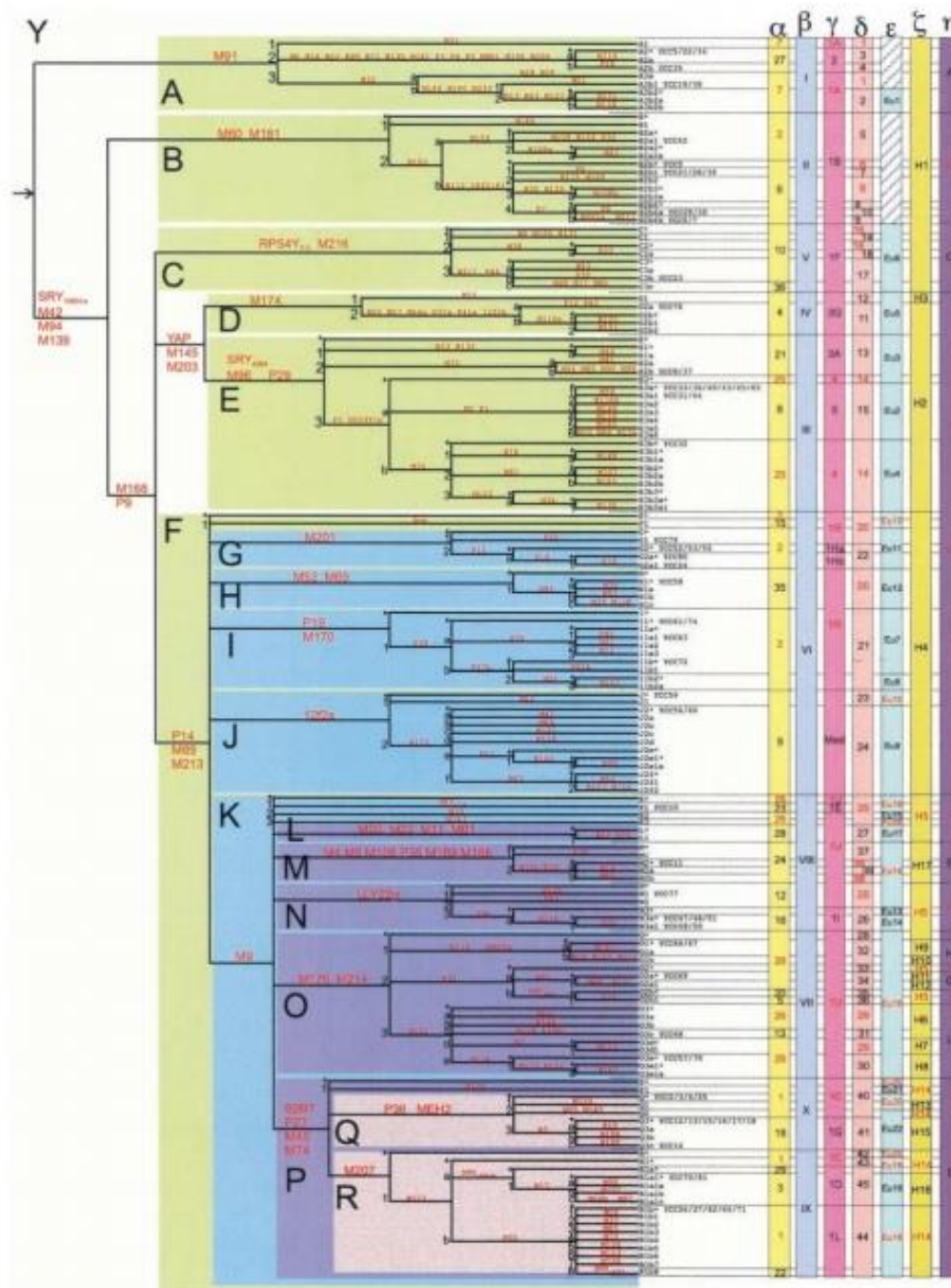


Figura 17 - Árvore dos haplogrupos binários do cromossoma Y proposta pelo Y Chromosome Consortium. Adaptado de Y Chromosome Consortium, 2002.

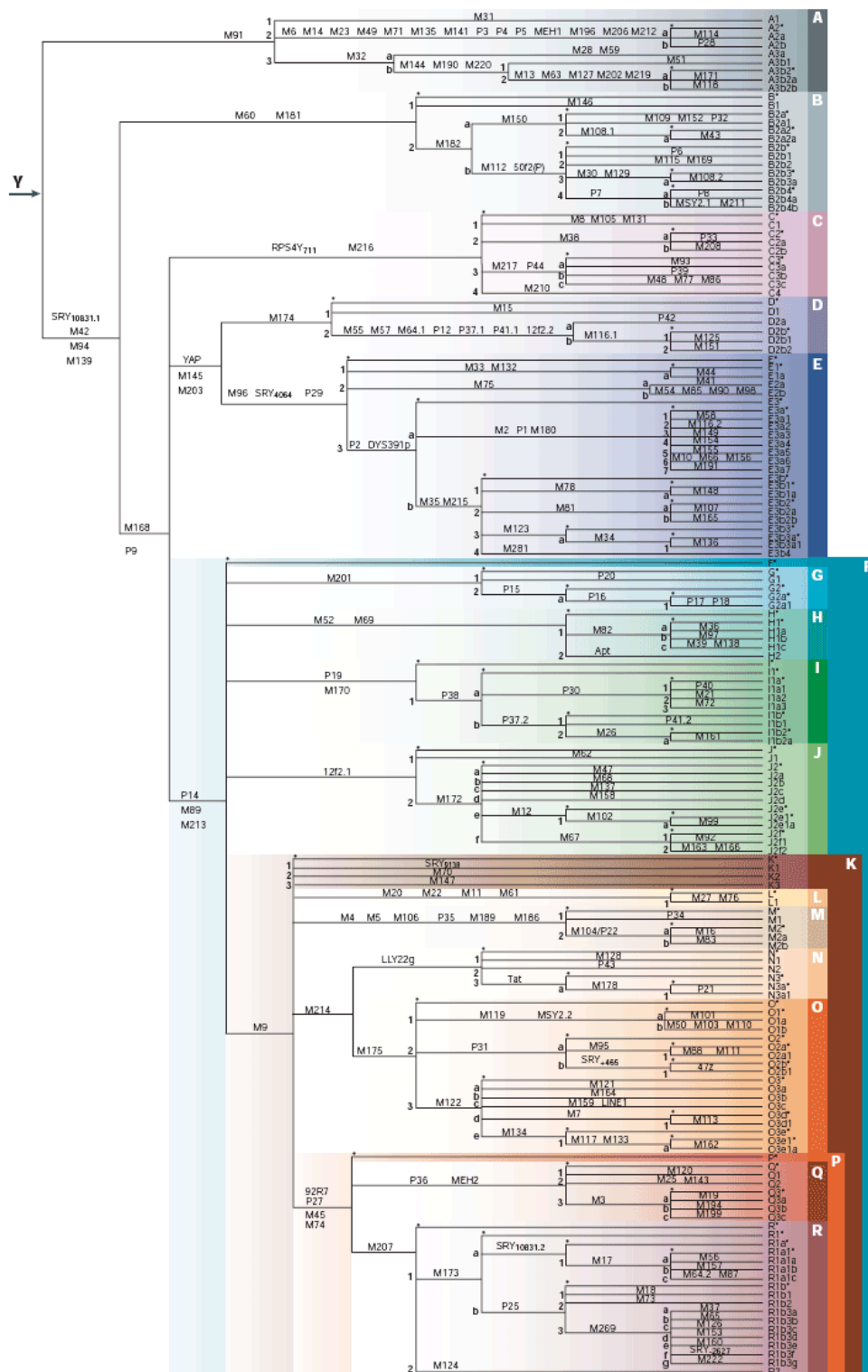


Figura 18 - Árvore dos haplogrupos binários do cromossoma Y proposta pelo YCC e revista por Karafet *et al.* (2008). Adaptado de Karafet *et al.*, 2008.

3.1 Análise do Multiplex: Estudo dos Polimorfismos

O estudo de SNPs do Cromossoma Y é de grande utilidade para obter respostas acerca dos padrões de migração do homem moderno. Essa informação é dada através dos polimorfismos da linhagem paterna, o que permitiu, pela distribuição dos haplogrupos, a construção da Árvore de Haplogrupos Binários do Cromossoma Y (Carvalho, 2010; Jobling e Tyler-Smith, 2003).

A informação genética presente neste cromossoma é transmitida integralmente de pai para filho, sem sofrer recombinação e com baixa taxa de mutação, permitindo que os SNPs sejam considerados uma ferramenta essencial para o estudo da origem e expansão populacional (Butler, 2012; Pinheiro, 2010).

Para esta investigação, diversos Y-SNPs foram selecionados numa reação *multiplex*, de modo a que a caracterização das duas populações fosse correta. A escolha deste multiplex com 9 Y-SNPs: SRY1532, M9, M70, M213, M22, Tat, 92R7, M173 e P25, permite a identificação das grandes clades da Árvore de Haplogrupos Binários do Cromossoma Y, e, conforme literatura já publicada, representam os haplogrupos com uma maior presença na Europa (Brion *et al.*, 2004).

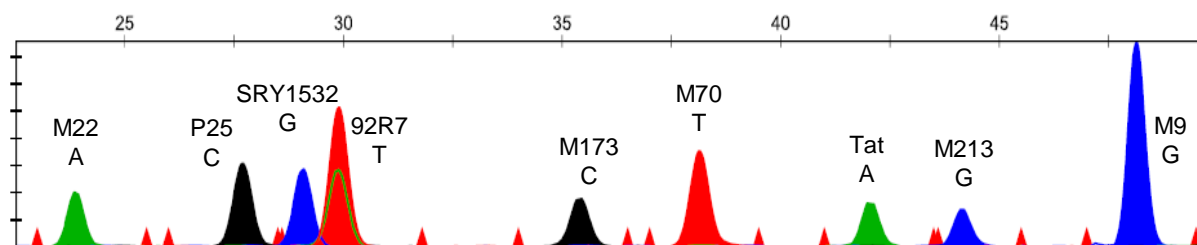


Figura 19 - Eletróforegrama dos 9 Y-SNPs que se apresentam sob a forma de fragmentos com diferentes tamanhos e cores, conforme a sua variação nucleotídica.

Os SNPs, sendo marcadores menos discriminativos do que os STRs, terão de ser analisados em maior número para se obter uma correta caracterização das populações em estudo (Carvalho, 2010). Este *multiplex* é o primeiro passo para a determinação dos haplogrupos, aconselhando-se posteriormente uma nova escolha de Y-SNPs, mais específicos de determinados grupos populacionais para classificar subclades da Árvore de haplogrupos binários do Cromossoma Y.

A caracterização dos haplogrupos do Cromossoma Y de cada indivíduo das duas populações foi efetuada por comparação com a Árvore de haplogrupos binários do Cromossoma Y (Jobling e Tyler-Smith, 2003; Karafet *et al.*, 2008; The Y Chromosome Consortium, 2002) e encontra-se mais detalhada no anexo VI – tabelas 14 e 15.

Com a análise dos 150 indivíduos das duas populações, foram identificados três diferentes haplogrupos em cada população, na caucasiana e cabo-verdiana: haplogrupo R1* (Fig. 20), F (xK) (Fig.21) e haplogrupos B,C,D ou E (Fig.22).

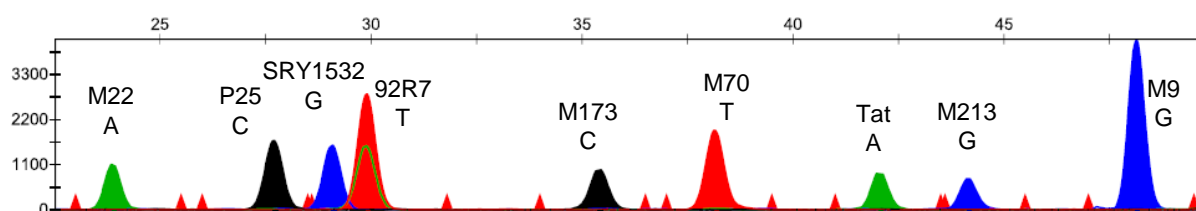


Figura 20 - Eletroforegrama característico do haplogrupo R1*.

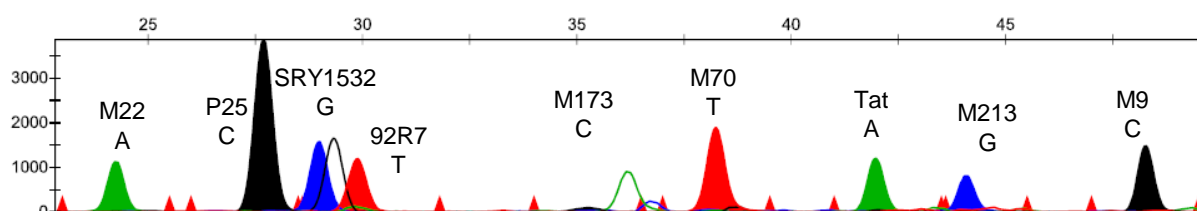


Figura 21 - Eletroforegrama característico do haplogrupo F (xK).

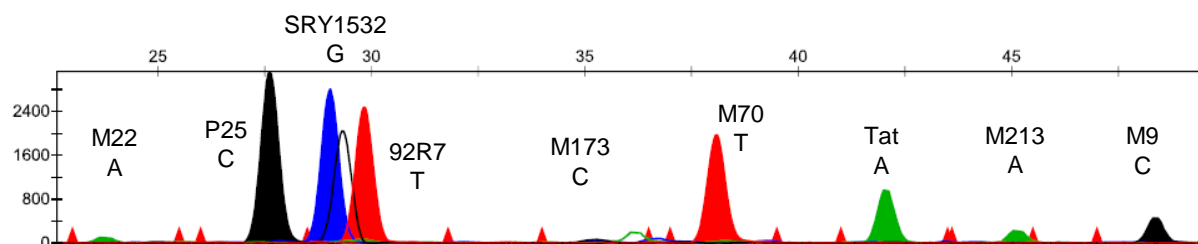


Figura 22 - Eletroforegrama característico dos haplogrupos B,C,D ou E.

3.2 Cálculo das frequências: Estudo dos Haplogrupos

O estudo dos haplogrupos foi realizado após a caracterização dos 9 Y-SNPs presentes no *multiplex*, encontrando-se descrito com mais detalhes no anexo VI, tabelas 14 e 15.

Foram definidos pelo menos três haplogrupos para os caucasianos e caboverdianos segundo a nomenclatura do YCC (The Y Chromosome Consortium, 2002), revista por *Jobling e Tyler-Smith* (Jobling e Tyler-Smith, 2003) e por *Karafet et al.* (Karafet *et al.*, 2008).

As frequências de cada haplogrupo de cada população foram determinadas com o programa de organização de dados IBM-SPSS versão 21.0, encontrando-se descrito nas tabelas 12 e 13.

Tabela 12 - Frequências dos haplogrupos presentes na população caucasiana, especificamente portuguesa.

Caucasianos		
Haplogrupo	Frequência	Porcentagem
R1*	52	69%
F (xK)	17	23%
B,C,D ou E	6	8%

Tabela 13 - Frequências dos haplogrupos presentes na população cabo-verdiana

Cabo Verde		
Haplogrupo	Frequência	Porcentagem
R1*	48	64%
F (xK)	9	12%
BCD ou E	18	24%

Os 9 Y-SNPs selecionados para este estudo estão associados à caracterização da população europeia (Brion *et al.*, 2004). Como consequência alguns indivíduos, nomeadamente os pertencentes aos haplogrupos B,C,D ou E, ficaram com a sua caracterização aquém do que poderia ter sido realizado. A solução seria introduzir na análise novos SNPs do cromossoma Y que permitem identificar qual o haplogrupo pertencente a uma destas Clades.

Ambas as populações são representadas pelos mesmos haplogrupos, R1*, F (xK) e haplogrupos B,C,D ou E, mas com frequências ligeiramente diferentes.

HAPLOGRUPOS DA POPULAÇÃO CAUCASIANA

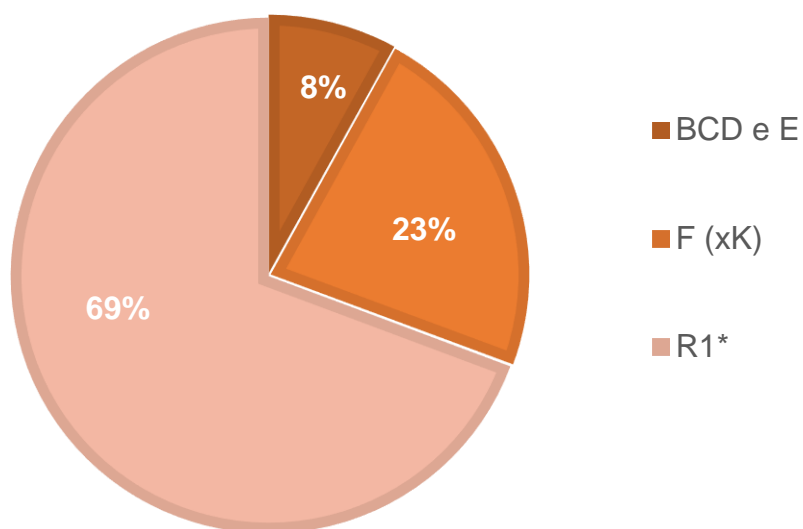


Figura 23 - Representação gráfica dos haplogrupos da população caucasiana

HAPLOGRUPOS DA POPULAÇÃO CABO-VERDIANA

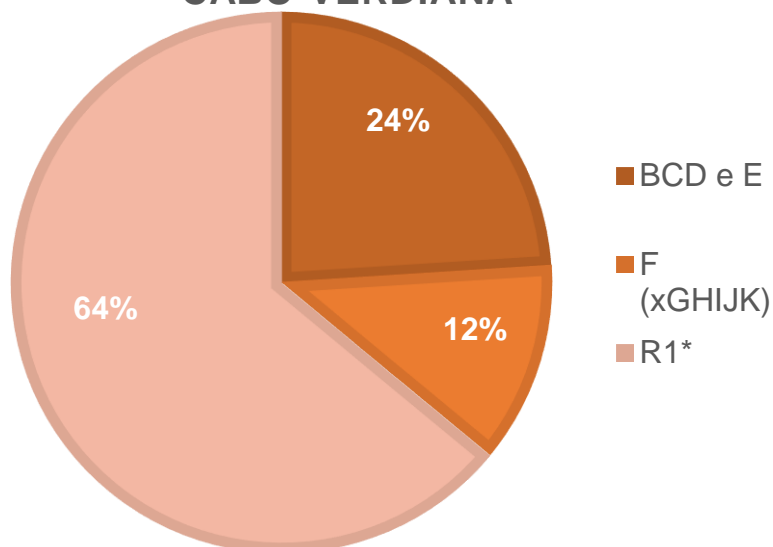


Figura 24 - Representação gráfica dos haplogrupos da população cabo-verdiana

Os haplogrupos B,C,D ou E, na população caucasiana representam 8% e na população cabo-verdiana representam cerca de 24%, têm em comum a ausência de mutação no M213. O haplogrupo B, restrito à África Subsariana, é proposto como a origem do homem moderno devido à sua ancestralidade e a sua restrição ao local de origem. Os haplogrupos C e D estão associados ao continente asiático. O C é bastante comum no Japão e Oceânia e o D na Ásia Central. O haplogrupo E estende-se a todo o continente africano, admitindo-se que foi a origem da expansão do homem moderno para fora de África, estando também presente fora deste (Chiaroni, Underhill e Cavalli-Sforza, 2009; Karafet *et al.*, 2008; Onofri *et al.*, 2006; Underhill e Kivisild, 2007).

O haplogrupo F (xK) representado por 23 % da população caucasiana e 12% na população de Cabo Verde caracteriza-se pela inexistência de mutação em M9. Este haplogrupo é resultante da migração inicial do homem moderno para fora do continente africano, sendo uma linhagem muito regular na Ásia Meridional, Sudoeste Africano e Médio Oriente (Chiaroni, Underhill e Cavalli-Sforza, 2009; Jobling e Tyler-Smith, 2003; Karafet *et al.*, 2008; Underhill e Kivisild, 2007).

O haplogrupo R1*, determinado pela ausência de mutação no P25, está presente em 69% da população caucasiana e 64% da cabo-verdiana. É um haplogrupo tipicamente europeu, estando presente na maioria desta população contudo, apresenta uma forte presença em todo o mundo (Chiaroni, Underhill e Cavalli-Sforza, 2009; Karafet *et al.*, 2008; Underhill e Kivisild, 2007).

A análise comparativa entre as duas populações, e após a interpretação dos haplogrupos por comparação com a árvore dos haplogrupos binários do cromossoma Y proposta pelo YCC, permitiu concluir que os resultados obtidos estão

em concordância com estudos anteriores bem como com os dados históricos (Gonçalves *et al.*, 2003; Grassi, 2006; Lopes, 2011).

Na população caucasiana, de origem portuguesa, o haplogrupo com uma maior representação é o R1* que tem origem tipicamente europeia. Na população cabo-verdiana podemos observar o mesmo, o que pode resultar de factos históricos, visto que esta ilha terá sido colonizada por portugueses na época dos descobrimentos, resultando em cruzamentos com as mulheres escravas africanas que povoavam Cabo Verde. Tendo sido mais tarde povoada por outros países europeus, o que contribui para a elevada percentagem deste haplogrupo (Carling e Batalha, 2008; Organização Internacional para as Migrações, 2009).

Relativamente aos restantes haplogrupos, o F (xK) apresenta uma maior percentagem na população caucasiana. Na população cabo-verdiana, a segunda maior representação pertence aos haplogrupos B,C,D ou E. As condições de emigração referidas na introdução levaram a um cruzamento entre homens colonizadores e mulheres africanas, justificando-se a maior percentagem dos haplogrupos B,C,D ou E na população cabo-verdiana (Brehm *et al.*, 2002; Grassi, 2006).

A presença do haplogrupo F (xK) em maior abundância na população caucasiana será reflexo da evolução populacional, resultando da expansão do homem justificada pela teoria *Out of África* (Brion *et al.*, 2004; Muro *et al.*, 2011; Semino *et al.*, 2000; Sims, Garvey e Ballantyne, 2007).

O estudo mais aprofundado de ambas as populações necessita da aplicação de novos Y-SNPs mais específicos para estes haplogrupos, de acordo com a Árvore

dos Haplogrupos Binários do Cromossoma Y, divididos em diversos *multiplexes* e proceder novamente à sua análise.

3.3 Distância Genética

Os Y-SNPs são essenciais para estudos filogenéticos a fim determinar a história das migrações (Carvalho, 2010).

A análise dos haplogrupos de uma população, a sua frequência, assim como a sua posição na árvore dos haplogrupos do cromossoma Y são determinantes para a construção de bases de dados populacionais (Karafet *et al.*, 2008; The Y Chromosome Consortium, 2002).

Esta investigação dos haplogrupos do Cromossoma Y, determinados como polimorfismos binários, é a forma mais comum do estudo da expansão da população assim como a existente variabilidade genética entre elas (Pinheiro, 2010).

Após a determinação dos polimorfismos binários na população caucasiana e na cabo-verdiana, baseando-se numa comparação com a árvore de haplogrupos proposta pelo YCC (The Y Chromosome Consortium, 2002), corrigida por *Jobling e Tyler-Smith* (Jobling e Tyler-Smith, 2003) e por *Karafet et al.* (Karafet *et al.*, 2008), procedeu-se à classificação dos haplogrupos encontrados em ambas as populações.

Posteriormente, prosseguiu-se ao cálculo da distância genética. A significância de distância genética entre as populações correspondentes vai determinar a variação genética intra e interpopulacional. A distância genética calculada através da *F de Wright* (1951) determinou o valor F_{ST} igual a 0,0267. O valor obtido representa a probabilidade de dois indivíduos, pertencentes a duas populações diferentes, possuírem o mesmo polimorfismo por descendência (Excoffier e Lischer, 2010).

A significância deste valor é verificada por um teste estatístico apropriado. O p_{Value} , obtido foi 0.0545 ($p > 0.05, \alpha = 0.05$), assim, o valor da distância genética obtida não é estatisticamente significativo. Conclui-se que não existem evidências estatísticas de distância genética entre as duas populações (Excoffier e Lischer, 2010).

Este resultado, mais uma vez, encontra-se em conformidade com os dados históricos. Existe uma proximidade genética entre estas as duas populações, provavelmente consequência dos cruzamentos entre os dois povos que advém já da época dos descobrimentos e posterior colonização por parte dos portugueses e mais tarde dos restantes europeus (Organização Internacional para as Migrações, 2009).

Todavia, a inexistente evidência de distância genética pode também ser justificada pela elevada presença da população de Cabo Verde em Portugal, sendo a segunda maior comunidade de emigrantes no nosso país.

Ainda assim, como o valor da estatística de teste está muito próximo do nível de significância, a introdução de mais Y-SNPs na análise poderia contribuir para diferenciar as duas populações e assim, deste modo, aumentar a distância genética entre as duas populações.

CONCLUSÕES

A análise dos polimorfismos binários é uma abordagem comum para o estudo da origem e expansão do homem moderno. Através da análise em simultâneo de *Single Nucleotide Polymorphisms* do cromossoma Y num *multiplex* é possível obter uma resposta rápida e eficaz relativamente à ascendência do indivíduo em estudo.

Apesar de todos os avanços que se têm realizado com os Y-SNPs, estes continuam a ser um auxílio dos *Short Tandem Repeat* na prática forense, sendo de grande utilidade quando os resultados destes não são conclusivos.

As vantagens de aplicação dos Y-SNPs podem verificar-se em situações de acidentes de massa ou quando o corpo está em avançado estado de composição, pois a identificação pessoal fica comprometida. Nestas situações pode recorrer-se à aplicação de SNPs pois para estudar estes marcadores são necessários fragmentos ADN com cerca de 100 pb enquanto o estudo de muitos STRs requer fragmentos de ADN com cerca de 300 a 400 pb. Outras situações de interesse forense poderão ser alguns processos criminais. A aplicação dos Y-SNPs nestas circunstâncias poderá contribuir para determinar qual o grupo étnico pertencente ao criminoso, contribuindo para uma nova orientação à investigação.

Um exemplo desta cooperação terá ocorrido durante os atentados de 11 de Março em Madrid. Com a explosão dos engenhos nos comboios, os corpos ficaram bastante degradados dificultando a identificação das vítimas assim como dos terroristas. Procedeu-se então à análise de uma seleção de SNPs a fim de se diferenciar a ascendência europeia e norte africana. Conseguiu-se chegar a um resultado: uma das amostras num engenho que não tinha explodido era de um

indivíduo de origem argelina. Ressalva-se a importância da adição de SNPs com uma elevada especificidade, pois ao associar o polimorfismo a uma subclade, aumentou substancialmente a eficácia da investigação (Phillips *et al.*, 2009).

Com a realização deste estudo, confirmou-se que a ancestralidade das linhagens paternas das duas populações está de acordo com estudos já publicados e em conformidade com os dados históricos.

A evidência de não existirem distâncias genéticas significativas com o conjunto de Y-SNPs estudados confirma a colonização dos portugueses nesta ilha e o seu cruzamento com os escravos africanos que habitavam Cabo Verde.

Nesta investigação, a escolha recaiu sobre 75 homens cuja de origem é cabo-verdiana, a segunda maior comunidade estrangeira presente no nosso país, e 75 homens de origem portuguesa. O propósito seria uma caracterização destas duas populações, com vista a estes resultados lograrem nos laboratórios forenses em perícias médico-legais, nomeadamente casos de investigação de parentesco ou em cenas de crime.

Para a descrição destas duas populações ser mais completa, sugere-se a aplicação de SNPs, mais específicos dos haplogrupos menos explorados.

Com a finalidade de os Y-SNPs entrarem na rotina dos laboratórios forenses são necessários mais estudos, mais trabalhos inter-laboratoriais com o desígnio final de construir uma melhor base de dados com resultados estatísticos em concordância.

REFERÊNCIAS BIBLIOGRÁFICAS

ADAMS, S. *et al.* - The case of the unreliable SNP: recurrent back-mutation of Y-chromosomal marker P25 through gene conversion. **Forensic Science International**. 159:1 (2006) 14–20. doi: 10.1016/j.forsciint.2005.06.003.

ALONSO, A. *et al.* - Specific quantification of human genomes from low copy number DNA samples in forensic and ancient DNA studies. **Croatian Medical Journal**. 44:3 (2003) 273–280.

AMORIM, A.; PEREIRA, L. - Pros and cons in the use of SNPs in forensic kinship investigation: A comparative analysis with STRs. **Forensic Science International**. 150:1 (2005) 17–21. doi: 10.1016/j.forsciint.2004.06.018.

ANDRÉASSON, H.; GYLLENSTEN, U.; ALLEN, M. - Real-time DNA quantification of nuclear and mitochondrial DNA in forensic analysis. **BioTechniques**. 33:2 (2002) 402–411.

APPLIED BIOSYSTEMS - SNaPshot Kit Analysis Getting Started Guide (GeneMapper Software Version 4.1). **Applied Biosystems**. (2009).

APPLIED BIOSYSTEMS - ABI Prism® SNaPshot™ Multiplex Kit Protocol. **Applied Biosystems**. (2010) 1 – 42.

APPLIED BIOSYSTEMS - Quantifiler® Kits- User Manual. **Applied Biosystems**. (2012).

APPLIED BIOSYSTEMS - SNaPshot® Multiplex System for SNP genotyping One system , many applications. **Applied Biosystems**. (2012) 1–4.

BELL, J. R. - A Simple Way to Treat PCR Products Prior to Sequencing Using ExoSAP-IT®. **BioTechniques**. 44:6 (2008) 834. doi: 10.2144/000112890.

BIO-RAD LABORATORIES - Chelex 100 and Chelex 20 Chelating Ion Exchange Resin Instruction Manual. California, USA. (2000) 1–24.

BOUAKAZE, C. *et al.* - First successful assay of Y-SNP typing by SNaPshot minisequencing on ancient DNA. **International Journal of Legal Medicine**. 121 (2007) 493–499. doi: 10.1007/s00414-007-0177-3.

BREHM, A. *et al.* - Mitochondrial portrait of the Cabo Verde archipelago: the Senegambian outpost of Atlantic slave trade. **Annals of Human Genetics**. 66:Pt 1 (2002) 49–60. doi:10.1017/S0003480001001002.

BRION, M. *et al.* - Hierarchical analysis of 30 Y-chromosome SNPs in European populations. **International Journal of Legal Medicine**. 119 (2004) 10–15. doi: 10.1007/s00414-004-0439-2.

BRION, M. *et al.* - 29 Y-chromosome SNP analysis in European populations. **International Congress Series**. 1261 (2004) 73–75. doi: 10.1016/S0531-5131(03)01822-3.

BRION, M. *et al.* - A collaborative study of the EDNAP group regarding Y-chromosome binary polymorphism analysis. **Forensic Science International**. 153 (2005) 103–108. doi: 10.1016/j.forsciint.2005.06.002.

BRIÓN, M. - Y Chromosome SNP Analysis Using the Single-Base Extension: A Hierarchical Multiplex Design. Em CARRACEDO, A. (Ed.) - **Methods in Molecular Biology: Forensic DNA Typing Protocols**. Totowa, NJ: Humana Press Inc., 2005 v. 297. p. 229–242.

BRIÓN, M. *et al.* - Introduction of an single nucleotide polymorphism-based «Major Y-chromosome haplogroup typing kit» suitable for predicting the geographical origin of male lineages. **Electrophoresis**. 26 (2005) 4411–4420. doi: 10.1002/elps.200500293.

BRIÓN, M. *et al.* - Analysis of 29 Y-chromosome SNPs in a single multiplex useful to predict the geographic origin of male lineages. **International Congress Series**. 1288 (2006) 13–15. doi: 10.1016/j.ics.2005.09.140.

BRISIGHELLI, F. *et al.* - Y-chromosomal and mitochondrial markers: A comparison between four population groups of Italy. **International Congress Series**. 1288 (2006) 91–93. doi: 10.1016/j.ics.2005.10.061.

BROOKES, A. J. - The essence of SNPs. **GENE**. 234:2 (1999) 177–186. doi: 10.1016/S0378-1119(99)00219-X.

BRUMFIELD, R. *et al.* - The utility of single nucleotide polymorphisms in inferences of population history. **Trends in Ecology and Evolution**. 18:5 (2003) 249–256. doi: 10.1016/S0169-5347(03)00018-1.

BUDOWLE, B.; DAAL, A. V. - Forensically relevant SNP classes. **BioTechniques**. 44:5 (2008) 603–610. doi: 10.2144/000112806.

BUTLER, J. M. - DNA Biology Review. Em BUTLER, J. M. (Ed.) - **Forensic DNA typing: biology, technology, and genetics of STR markers**. 2ª edição. USA: Elsevier, 2005. p. 17–31.

BUTLER, J. M. - Y Chromosome DNA testing. Em BUTLER, J. M. (Ed.) - **Forensic DNA typing: biology, technology, and genetics of STR markers**. 2ª edição. USA: Elsevier, 2005. p. 201–240.

BUTLER, J. M. - Overview and History of DNA typing. Em BUTLER, J. M. (Ed.) - **Forensic DNA typing: biology, technology, and genetics of STR markers**. 2ª edição. USA: Elsevier, 2005. p. 1–13.

BUTLER, J. M. - Single Nucleotide Polymorphisms and other bi-allelic markers. Em BUTLER, J. M. (Ed.) - **Forensic DNA typing: biology, technology, and genetics of STR markers**. 2ª edição. USA: Elsevier, 2005. p. 181–200.

BUTLER, J. M. - Sample Collection, DNA Extraction and Quantification. Em BUTLER, J. M. (Ed.) - **Forensic DNA typing: biology, technology, and genetics of STR markers**. 2ª edição. USA: Elsevier, 2005. p. 33–62.

BUTLER, J. M. - The Polymerase Chain Reaction (DNA Amplification). Em BUTLER, J. M. (Ed.) - **Forensic DNA typing: biology, technology, and genetics of STR markers**. 2ª edição. USA: Elsevier, 2005. p. 63–83.

BUTLER, J. M. - Basics of DNA Biology and Genetics. Em BUTLER, J. M. (Ed.) - **Fundamentals of forensic DNA typing**. California, USA: Academic Press, 2010. p. 19–41.

BUTLER, J. M. - DNA Extraction. Em BUTLER, J. M. (Ed.) - **Fundamentals of forensic DNA typing**. California, USA: Academic Press, 2010. p. 99–110.

BUTLER, J. M. - DNA Quantification. Em BUTLER, J. M. (Ed.) - **Fundamentals of forensic DNA typing**. California, USA: Academic Press, 2010. p. 111–124.

BUTLER, J. M. - DNA Amplification (The polymerase chain reaction). Em BUTLER, J. M. (Ed.) - **Fundamentals of forensic DNA typing**. California, USA: Academic Press, 2010. p. 125–146.

BUTLER, J. M. - Y-Chromosome DNA Testing. Em BUTLER, J. M. (Ed.) - **Advanced Topics in Forensic DNA Typing**. California, USA: Academic Press, 2012. p. 371–404.

BUTLER, J. M. - DNA Extraction Methods. Em BUTLER, J. M. (Ed.) - **Advanced Topics in Forensic DNA Typing: Methodology**. California, USA: Academic Press, 2012. ISBN 9780123745132. p. 29–48.

BUTLER, J. M. - Single Nucleotide Polymorphisms and Applications. Em BUTLER, J. M. (Ed.) - **Advanced Topics in Forensic DNA Typing**. California, USA: Academic Press, 2012. p. 347–370.

BUTLER, J. M. - DNA Quantification. Em BUTLER, J. M. (Ed.) - **Advanced Topics in Forensic DNA Typing**. California, USA: Academic Press, 2012. p. 49–68.

BUTLER, J. M. - PCR Amplification: Capabilities and Cautions. Em BUTLER, J. M. (Ed.) - **Advanced Topics in Forensic DNA Typing**. California, USA: Academic Press, 2012. p. 69–98.

BUTLER, J. M.; COBLE, M. D.; VALLONE, P. M. - STRs vs. SNPs: Thoughts on the future of forensic DNA testing. **Forensic Science, Medicine, and Pathology**. 3 (2007) 200–205. doi: 10.1007/s12024-007-0018-1.

CARLING, J.; BATALHA, L. - Cape Verdean: Migration and Diaspora. Em BATALHA, L.; CARLING, J. (Eds.) - **Transnational Archipelago: Perspectives on Cape Verdean Migration and Diaspora**. Amsterdão: Amsterdam University Press, 2008.

CARRACEDO, A.; SÁNCHEZ-DIZ, P. - Forensic DNA-Typing Technologies. Em CARRACEDO, A. (Ed.) - **Methods in Molecular Biology: Forensic DNA Typing Protocols**. Totowa, NJ: Humana Press Inc., 2005. v. 297. p. 1–11.

CARRACEDO, A. - Forensic DNA Typing Protocols. **Methods in Molecular Biology**. Totowa, NJ. 297 (2005). doi: 10.1385/1592598676.

CARVALHO, M. R. - Single nucleotide polymorphisms (SNPs) com aplicação forense. Em PINHEIRO, M. F. (Ed.) - **Genética forense: perspectivas da identificação genética**. Porto: Edições Universidade Fernando Pessoa, 2010. p. 187–198.

CHAKRABORTY, R. *et al.* - The utility of short tandem repeat loci beyond human identification: Implications for development of new DNA typing systems. **Electrophoresis**. 20:8 (1999) 1682–1696. doi: 10.1002/(SICI)1522-2683(19990101)20:8<1682::AID-ELPS1682>3.0.CO;2-Z.

CHARLESWORTH, B. - The Evolution of Sex Chromosomes. **Science**. 251 (1991) 1030–1033.

CHIARONI, J.; UNDERHILL, P. A.; CAVALLI-SFORZA, L. L. - Y chromosome diversity, human expansion, drift, and cultural evolution. **Proceedings of the National Academy of Sciences of the United States of America**. 106:48 (2009) 20174–20179. doi: 10.1073/pnas.1008738107.

CRUCIANI, F. *et al.* - Human Y chromosome haplogroup R-V88: a paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. **European Journal of Human Genetics**. 18:7 (2010) 800–807. doi: 10.1038/ejhg.2009.231.

DANIEL, R.; WALSH, S. J.; PIPER, A. - Investigation of single-nucleotide polymorphisms associated with ethnicity. **International Congress Series**. 1288 (2006) 79–81. doi: 10.1016/j.ics.2005.11.131.

DUARTE, C. *et al.* - The early Upper Paleolithic human skeleton from the Abrigo do Lagar Velho (Portugal) and modern human emergence in Iberia. **Proceedings of the National Academy of Sciences of the United States of America**. 96:13 (1999) 7604–7609. doi: 10.1073/pnas.96.13.7604.

EXCOFFIER, L.; LISCHER, H. - Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**. 10:3 (2010) 564–567. doi: 10.1111/j.1755-0998.2010.02847.x.

FERNANDES, A. T. *et al.* - Genetic differentiation of the Cabo Verde Archipelago population analysed by STR polymorphisms. **Annals of Human Genetics**. 67:4 (2003) 340–347. doi: 10.1046/j.1469-1809.2003.00029.x.

FONDEVILA, M. *et al.* - Forensic performance of insertion–deletion marker systems. **Forensic Science International: Genetics Supplement Series**. 3 (2011) 443–444. doi: 10.1186/1471-2156-9-8.

GEPPERT, M.; ROWER, L. - SNaPshot® Minisequencing Analysis of Multiple Ancestry- Informative Y-SNPs Using Capillary Electrophoresis. Em ALONSO, ANTONIO (Ed.) - **DNA Electrophoresis Protocols for Forensic Genetics, Methods in Molecular Biology**. Totowa, NJ: Humana Press Inc., 2012. v. 830. p. 127–140.

GILL, P. - An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. **International Journal of Legal Medicine**. 114 (2001) 204–210. doi: 10.1007/s004149900117.

GÓIS, P. - **A Emigração Cabo-Verdiana para (e na) Europa e a sua inserção em mercados de trabalho locais: Lisboa, Milão, Roterdão**. Lisboa: Alto-Comissariado para a Imigração e Minorias Étnicas. 2002.

GONÇALVES, R. *et al.* - Y-chromosomes lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers. **Human Genetics**. 113 (2003) 467–472. doi: 10.1007/s00439-003-1007-4.

GONÇALVES, R. *et al.* - Y-chromosome lineages from Portugal, Madeira and Açores record elements of Sephardim and Berber ancestry. **Annals of Human Genetics**. 69 (2005) 443–454. doi: 10.1016/j.ics.2005.09.099.

GOODWIN, W.; LINACRE, A.; HADI, S.- DNA struture and the genome. Em GOODWIN, W.; LINACRE, A.; HADI, S. (Eds.)- **An Introduction to Forensic Genetics**. 2ª edição. Inglaterra: John Willey & Sons, Ltd, 2011. p. 11–20.

GOODWIN, W.; LINACRE, A.; HADI, S.- Single nucleotide polymorphisms. Em GOODWIN, W.; LINACRE, A.; HADI, S. (Eds.)- **An Introduction to Forensic Genetics**. 2ª edição. Inglaterra: John Willey & Sons, Ltd, 2011. p. 145–154.

GOODWIN, W.; LINACRE, A.; HADI, S.- DNA Extraction and quantification. Em GOODWIN, W.; LINACRE, A.; HADI, S. (Eds.)- **An Introduction to Forensic Genetics**. 2ª edição. Inglaterra: John Willey & Sons, Ltd, 2011. p. 37–52.

GOODWIN, W.; LINACRE, A.; HADI, S. - Polymerase Chain Reaction. Em GOODWIN, W.; LINACRE, A.; HADI, S. (Eds.) - **An Introduction to Forensic Genetics**. 2ª edição. Inglaterra: John Willey & Sons, Ltd, 2011. p. 53–66.

GRASSI, M. - **Cabo Verde pelo Mundo: O Género e a Diáspora Cabo-Verdiana**. Lisboa: Instituto de Ciências Sociais - Universidade de Lisboa, 2006

GRILO, L. T. - O DNA. Em VIDEIRA, A. (Ed.) - **Engenharia genética - Princípios e Aplicações**. Lisboa: Lidel - edições técnicas, lda, 2001. p. 3–14.

GUSMÃO, L. *et al.* - Y Chromosome specific polymorphisms in forensic analysis. **Legal Medicine**. 1 (1999) 55–60. doi: 10.1016/S1344-6223(99)80013-3.

GUSMÃO, L.; ALVES, C. - Marcadores dos cromossomas sexuais. Revelância na identificação genética. Em PINHEIRO, M. F. (Ed.) - **Genética forense: perspectivas da identificação genética**. Porto: Edições Universidade Fernando Pessoa, 2010. p. 103–138.

HEID, C. A. *et al.* - Real Time Quantitative PCR. **Genome Research**. 6 (1996) 986–994. doi: 10.1101/gr.6.10.986.

HIGUCHI, R. *et al.* - Kinetic PCR Analysis: Real-Time Monitoring of DNA Amplification Reactions. **Bio/Technology**. 11:9 (1993) 1026–1030. doi: 10.1038/nbt0993-1026.

HUGHES, J. F. *et al.* - Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee. **Nature**. 437:7055 (2005) 100–103. doi: 10.1038/nature04776.

INSTITUTO NACIONAL DE ESTATÍSTICA - **Censos 2011: resultados definitivos - Portugal**. Lisboa: Instituto Nacional de Estatística. 2012.

INSTITUTO NACIONAL DE ESTATÍSTICA - **A População Estrangeira em Portugal**. Lisboa: Instituto Nacional de Estatística. 2012. 1–32.

INSTITUTO NACIONAL DE ESTATÍSTICA DE CABO VERDE - **IVº Recenseamento Geral Da População E De Habitação - Censo 2010**. Cabo Verde: Instituto Nacional de Estatística. 2010. 1–40.

JOBLING, M. A. - Y-chromosomal SNP haplotype diversity in forensic analysis. **Forensic Science International**. 118:2-3 (2001) 158–162. doi: 10.1016/S0379-0738(01)00385-1.

JOBLING, M. A.; PANDYA, A.; TYLER-SMITH, C. - The Y chromosome in forensic analysis and paternity testing. **International Journal of Legal Medicine**. 110:3 (1997) 118–124. doi: 10.1007/s004140050050.

JOBLING, M. A.; TYLER-SMITH, C. - Fathers and sons: the Y chromosome and human evolution. **Trends In Genetics**. 11:11 (1995) 449–456. doi: 10.1016/S0168-9525(00)89144-1.

JOBLING, M. A.; TYLER-SMITH, C. - New uses for new haplotypes: The human Y chromosome, disease and selection. **Trends in Genetics**. 16:8 (2000) 356–362. doi: 10.1016/S0168-9525(00)02057-6.

JOBLING, M. A.; TYLER-SMITH, C. - The human Y chromosome: An evolutionary marker comes of age. **Genetics**. 4:August (2003) 598–612. doi: 10.1038/nrg1124.

KARAFET, T. M. *et al.* - New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. **Genome Research**. 18 (2008) 830–838. doi: 10.1101/gr.7172008.

KOSOY, R. *et al.* - Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. **Human Mutation**. 30:1 (2009) 69–78. doi: 10.1002/humu.20822.

LAHN, B. T. *et al.* - The Human Y Chromosome, in the light of evolution. **Genetics**. 2 (2001) 207–216.

LAO, O. *et al.* - Proportioning Whole-Genome Single-Nucleotide-Polymorphism Diversity for the Identification of Geographic Population Structure and Genetic Ancestry. **American Journal of Human Genetics**. 78:4 (2006) 680–690. doi: 10.1086/501531.

LESSIG, R. *et al.* - SNP-genotyping on human Y-chromosome for forensic purposes: comparison of two different methods. **International Congress Series**. 1261 (2004) 334–336. doi: 10.1016/S0531-5131(03)01778-3.

LESSIG, R. *et al.* - Y-SNP-genotyping - a new approach in forensic analysis. **Forensic science international**. 154:2-3 (2005) 128–36. doi: 10.1016/j.forsciint.2004.09.129.

LI, L. *et al.* - SNP genotyping by multiplex amplification and microarrays assay for forensic application. **Forensic Science International**. 162 (2006) 74–79. doi: 10.1016/j.forsciint.2006.06.010.

LOPES, J. V. S. - **Distribuição de segmentos do cromossoma X com ancestralidade europeia e africana na população de cabo verde- implicações para o estudo da miscigenação em populações humanas e para a história do povoamento do arquipélago**. Porto: Universidade do Porto, 2011

MAKALOWSKI, W. - The human genome structure and organization. **Acta Biochimica Polonica**. 48:3 (2001) 587–598.

MARTINS, T.; CARVALHO, R. - Conceitos básicos de genética com aplicação á investigação forense. Em PINHEIRO, M. F. (Ed.) - **CSI Criminal**. Porto: Edições Universidade Fernando Pessoa, 2008. p. 81–92.

MATTOSO, J. - A formação da nacionalidade. Em TENGARRINHA, JOSÉ (Ed.) - **História de Portugal**. São Paulo, Brasil: Editora da Universidade do Sagrado Coração, 2000. p. 375.

MENGEL-JØRGENSEN, J. *et al.* - MALDI-TOF mass spectrometric detection of multiplex single base extended primers. A study of 17 Y-chromosome single-nucleotide polymorphisms. **Analytical Chemistry**. 76:20 (2004) 6039–6045. doi: 10.1021/ac049264k.

MULLIS, K. *et al.* - Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. **Cold Spring Harbor Symposia on Quantitative Biology**. 51 (1992) 263–273. doi: 10.1101/SQB.1986.051.01.032.

MURO, T. *et al.* - Simultaneous determination of seven informative Y chromosome SNPs to differentiate East Asian, European, and African populations. **Legal Medicine**. 13:3 (2011) 134–41. doi: 10.1016/j.legalmed.2011.01.001.

MYRES, N. M. *et al.* - A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. **European Journal of Human Genetics**. 19 (2011) 95–101. doi: 10.1038/ejhg.2010.146.

NICKLAS, J. A.; BUEL, E. - Quantification of DNA in forensic samples. **Analytical and Bioanalytical Chemistry**. 376 (2003) 1160–1167. doi: 10.1007/s00216-003-1924-z.

NIEVERGELT, C. M. *et al.* - Inference of human continental origin and admixture proportions using a highly discriminative ancestry informative 41-SNP panel. **Investigative Genetics**. 4:1 (2013) 1–16. doi: 10.1186/2041-2223-4-13.

NILSSON, M. *et al.* - Sensitive forensic analysis using the Pyrosequencing technology. **International Congress Series**. 1288 (2006) 625–627. doi: 10.1016/j.ics.2005.09.154.

NUSSBAUM, R.; MCINNES, R.; WILLARD, H. - The human genome: gene struture and function. Em NUSSBAUM, R.; MCINNES, R.; WILLARD, H. (Eds.) - **Thompson & Thompson Genetics in Medicine**. 8ª edição. Canadá: Elsevier, 2015. p. 21–41.

NUSSBAUM, R.; MCINNES, R.; WILLARD, H. - Introduction to the human genome. Em NUSSBAUM, R.; MCINNES, R.; WILLARD, H. (Eds.) - **Thompson & Thompson Genetics in Medicine**. 8ª edição. Canadá: Elsevier, 2015. p. 3–20.

NUSSBAUM, R.; MCINNES, R.; WILLARD, H. - Human Genetic Diversity: Mutation and Polymorphism. Em NUSSBAUM, R.; MCINNES, R.; WILLARD, H. (Eds.) -

Thompson & Thompson Genetics in Medicine. 8ª edição. Canadá: Elsevier, 2015. p. 43–55.

OLIVEIRA, I. T. DE - Emigração, retorno e reemigração na primeira metade do século XX. **Análise Social.** 42:184 (2007) 837–852.

ONOFRI, V. *et al.* - Development of multiplex PCRs for evolutionary and forensic applications of 37 human Y chromosome SNPs. **Forensic Science International.** 157:1 (2006) 23–35. doi: 10.1016/j.forsciint.2005.03.014.

ONORI, N. *et al.* - Post-mortem DNA damage: A comparative study of STRs and SNPs typing efficiency in simulated forensic samples. **International Congress Series.** 1288 (2006) 510–512. doi: 10.1016/j.ics.2005.09.139.

ORGANIZAÇÃO INTERNACIONAL PARA AS MIGRAÇÕES - **Migração em Cabo Verde, Perfil Nacional 2009.** Suíça: Organização Internacional para as Migrações. 2009.

OVEN, M. VAN *et al.* - Multiplex genotyping assays for fine-resolution subtyping of the major human Y-chromosome haplogroups E, G, I, J, and R in anthropological, genealogical, and forensic investigations. **Electrophoresis.** 34 (2013) 3029–3038. doi: 10.1002/elps.201300210.

OVEN, M. VAN; RALF, A.; KAYSER, M. - An efficient multiplex genotyping approach for detecting the major worldwide human Y-chromosome haplogroups. **International Journal of Legal Medicine.** 125 (2011) 879–885. doi: 10.1007/s00414-011-0605-2.

PATI, N. *et al.* - A comparison between SNaPshot, pyrosequencing, and biplex invader SNP genotyping methods: Accuracy, cost, and throughput. **Journal of Biochemical and Biophysical Methods.** 60 (2004) 1–12. doi: 10.1016/j.jbbm.2003.11.005.

PETKOVSKI, E. *et al.* - MALDI-TOF MS analysis of Y-SNPs in ancient samples. **International Congress Series.** 1288 (2006) 25–27. doi: 10.1016/j.ics.2005.12.008.

PHILLIPS, C. *et al.* - Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. **Forensic Science International: Genetics.** 1-4 (2007) 273–280. doi: 10.1016/j.fsigen.2007.06.008.

PHILLIPS, C. *et al.* - Ancestry analysis in the 11-M Madrid bomb attack investigation. **PLoS ONE.** 4:8 (2009) 1–10. doi: 10.1371/journal.pone.0006583.

PINHEIRO, M. F. - A perícia em genética e biologia forense - Criminalística forense. Em PINHEIRO, M. F. (Ed.) - **CSI Criminal.** Porto: Edições Universidade Fernando Pessoa, 2008. p. 7–38.

PINHEIRO, M. F. - Algumas perspectivas da identificação genética. Em PINHEIRO, M. F. (Ed.) - **Genética forense: perspectivas da identificação genética**. Porto: Edições Universidade Fernando Pessoa, 2010. p. 17–78.

POETSCH, M. *et al.* - Prediction of people's origin from degraded DNA - Presentation of SNP assays and calculation of probability. **International Journal of Legal Medicine**. 127 (2013) 347–357. doi: 10.1007/s00414-012-0728-0.

PULKER, H. *et al.* - Finding genes that underlie physical traits of forensic interest using genetic tools. **Forensic Science International: Genetics**. 1:2 (2007) 100–104. doi: 10.1016/j.fsigen.2007.02.009.

QIAGEN® - QIAGEN® Multiplex PCR Handbook. **Qiagen**®. (2010) 1–48.

QUINTÁNS, B. *et al.* - Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. **Forensic Science International**. 140:2-3 (2004) 251–257. doi: 10.1016/j.forsciint.2003.12.005.

REBELO DOS SANTOS, J. *et al.* - Imigrantes cabo-verdianos em Portugal: integração e sua percepção em relação aos portugueses. (2010) 3457–3482.

SANCHEZ, J. J. *et al.* - Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. **Forensic Science International**. 137:1 (2003) 74–84. doi: 10.1016/S0379-0738(03)00299-8.

SANCHEZ, J. J. *et al.* - A multiplex assay with 52 single nucleotide polymorphisms for human identification. **Electrophoresis**. 27:9 (2006) 1713–1724. doi: 10.1002/elps.200500671.

SANCHEZ, J. J.; BØRSTING, C.; MORLING, N. - Typing of Y chromosome SNPs with multiplex PCR methods. Em CARRACEDO, A. (Ed.) - **Methods in Molecular Biology: Forensic DNA Typing Protocols**. Totowa, NJ: Humana Press Inc., 2005. v. 297. p. 209–228.

SHELL, D. *et al.* - Multiplex typing with 5 Y-chromosomal SNPs. **International Congress Series**. 1288 (2006) 22–24. doi: 10.1016/j.ics.2005.12.011.

SCHULL, W. J. - Genetic structure of human populations. **Journal of toxicology and environmental health**. 5:1 (1979) 17–25.

SEMINO, O. *et al.* - The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. **Science**. 290 (2000) 1155–1159. doi: 10.1126/science.290.5494.1155.

SERDAR, T.; SEMA, D. - Y-SNP haplogroups in the Antalya population in Turkish Republic. **Romanian Society of Legal Medicine**. 17:1 (2009) 59–68.

SERVIÇO DE ESTRANGEIROS E FRONTEIRAS (SEF) - **Relatório de Imigração, Fronteiras e Asilo - 2013**. Lisboa, Portugal: Serviço de Estrangeiros e Fronteiras, 2014.

SIMS, L. M.; GARVEY, D.; BALLANTYNE, J. - Sub-populations within the major European and African derived haplogroups R1b3 and E3a are differentiated by previously phylogenetically undefined Y-SNPs. **Human Mutation**. 28 (2007) 1-7. doi: 10.1002/humu.9469.

SKALETISKY, H. *et al.* - The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. **Nature**. 423 (2003) 825–837. doi: 10.1038/nature01722.

SOBRINO, B.; BRIÓN, M.; CARRACEDO, A. - SNPs in forensic genetics: a review on SNP typing methodologies. **Forensic Science International**. 154:2-3 (2005) 181–94. doi: 10.1016/j.forsciint.2004.10.020.

SOBRINO, B.; CARRACEDO, A. - SNP typing in Forensic Genetics. Em CARRACEDO, A. (Ed.) - **Methods in Molecular Biology: Forensic DNA Typing Protocols**. Totowa, NJ: Humana Press Inc., 2005. v. 297. p. 107–126.

THE Y CHROMOSOME CONSORTIUM - A Nomenclature System for the Tree of Human Y-Chromosomal Binary Haplogroups. **Genome Research**. 12 (2002) 339–348. doi: 10.1101/gr.217602.polymorphisms.

TURNER, A. - SNPs on Chips: A New Source of Data for Y-Chromosome Studies. **Journal of Genetic Genealogy**. 4:1 (2008) 6–7.

UNDERHILL, P. A. *et al.* - Y chromosome sequence variation and the history of human populations. **Nature Genetics**. 26 (2000) 358–361. doi: 10.1038/81685.

UNDERHILL, P. A. *et al.* - The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. **Annals of Human Genetics**. 65 (2001) 43–62. doi: 10.1046/j.1469-1809.2001.6510043.x.

UNDERHILL, Peter A.; KIVISILD, Toomas - Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. **Annual review of genetics**. 41 (2007) 539–564. doi: 10.1146/annurev.genet.41.110306.130407.

VALVERDE, L. *et al.* - Improving the analysis of Y-SNP haplogroups by a single highly informative 16 SNP multiplex PCR-minisequencing assay. **Electrophoresis**. 34 (2013) 605–612. doi: 10.1002/elps.201200433.

VIDEIRA, A. - Reação em Cadeia da Polimerase (PCR). Em VIDEIRA, A. (Ed.) - **Engenharia genética - Princípios e Aplicações**. Lisboa: Lidel - edições técnicas, lda, 2001. p. 67–70.

WALSH, P. S.; METZGER, D. A.; HIGUCHI, R. - Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. **BioTechniques**. 10:4 (1991) 506–513.

WEI, W. *et al.* - A comparison of Y-chromosomal lineage dating using either resequencing or Y-SNP plus Y-STR genotyping. **Forensic Science International: Genetics**. 7:6 (2013) 568–572. doi: 10.1016/j.fsigen.2013.03.014.

WERLE, E. *et al.* - Convenient single-step, one tube purification of PCR products for direct sequencing. **Nucleic Acids Research**. 22:20 (1994) 4354–4355. doi: 10.1093/nar/22.20.4354.

WETTON, J. H.; TSANG, K. W.; KHAN, H. - Inferring the population of origin of DNA evidence within the UK by allele-specific hybridization of Y-SNPs. **Forensic science international**. 152:1 (2005) 45–53. doi: 10.1016/j.forsciint.2005.03.009.

WILLARD, H. F. - Tales of the Y chromosome. **Nature**. 423 (2003) 810–813. doi: 10.1038/423810a.

YANG, N. *et al.* - Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: Application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. **Human Genetics**. 118:3-4 (2005) 382–392. doi: 10.1007/s00439-005-0012-1.

ANEXOS

Anexo I – Protocolo utilizado para Extração de ADN a partir de manchas de sangue por Chelex® 100

1. Cada amostra de sangue, previamente colhida no cartão *Whatman™ FTA™* (GE Healthcare), é cortada um círculo com o auxílio de um alicate “*paper punch single hole*”. Este corte tem uma área de 3 mm² e coloca-se no fundo do tubo *ependorf* correspondente, já previamente esterilizado com radiações U.V. e identificado com o número da amostra;
2. Limpar sempre o alicate com álcool a 70% entre cada amostra, de modo a evitar as contaminações;
3. Adicionar 1 ml de água *Milli-Q* a cada tubo. Deixar a encubar 15 minutos a uma temperatura ambiente;
4. Agitar os tubos no *vórtex* durante 5 segundos, usando uma rotação baixa;
5. Centrifugar os tubos a 1400 rpm, numa centrífuga durante 3 minutos, a 20°C;
6. Cuidadosamente, retira-se com a micropipeta aproximadamente 940 µL do sobrenadante, de modo a não perturbar o *pellet*. Deve-se deixar cerca de 20 µL da solução com amostra em cada tubo;
7. Adicionar 180 µL de Chelex® 100 a 5% perfazendo um volume total de 200 µL. A solução de Chelex® 100 é homogeneizada antes de ser adicionada a cada amostra. ¹

¹ – Para esta operação deve-se cortar a ponta de cada pipeta para evitar a acumulação de resinas do Chelex®100 na ponta.

8. Incubar as amostras num banho de aquecimento a uma temperatura de 56°C durante 15 minutos;
9. Agitar no *vórtex* a alta velocidade durante 10 segundos;
10. Colocar os tubos em suportes flutuantes e de seguida, colocar estes no interior de copos de vidro com água a ferver. Contabiliza-se 8 minutos. Esta fase é necessária para o rompimento das células e consequente saída do ADN nuclear;
11. Agitar novamente as amostras no *vórtex*, aproximadamente 5 segundos a alta velocidade;
12. Centrifugar as amostras a 1400 rpm durante 5 minutos;
13. Armazenar as amostras no frigorífico a 4°C ou no congelador a -20°C, quando não se procede imediatamente à fase de amplificação do ADN das amostras.
14. Caso se congele as amostras, depois de descongeladas deve-se repetir o passo número 11 antes de se prosseguir.

Durante o processo de extração, deverá haver um tubo que servirá de controlo negativo. Esse tubo não contém qualquer tipo de amostra biológica, sendo constituído apenas por água e Chelex[®] 100. Este controlo é quantificado, amplificado e analisado para certificar que não há contaminação, sendo que não poderá obter qualquer resultado.

Anexo II – Protocolo utilizado para Quantificação de ADN com o Quantifiler® Human ADN Quantification Kit

1. Colocar a câmara de fluxo laminar a descontaminar com raios Ultravioleta.
Simultaneamente, ligar o *HID Real-Time PCR Analyse (ABI Prism® 7500)*;
2. Preparar a mistura do *ADN Standard*:

1	30 µL de <i>buffer</i> mais 10 µL de <i>ADN Standard</i>
2	30 µL de <i>buffer</i> mais 10 µL do tubo 1
3	30 µL de <i>buffer</i> mais 10 µL do tubo 2
4	30 µL de <i>buffer</i> mais 10 µL do tubo 3
5	30 µL de <i>buffer</i> mais 10 µL do tubo 4
6	30 µL de <i>buffer</i> mais 10 µL do tubo 5
7	30 µL de <i>buffer</i> mais 10 µL do tubo 6
8	30 µL de <i>buffer</i> mais 10 µL do tubo 7

Entre cada preparação do tubo com o *ADN Standard*, este deve ser agitado no *vórtex* durante 5 segundos;

Seguido aos primeiros 16 poços da placa com o *ADN Standard* (coloca-se em duplicado), segue-se o poço NTC que vai funcionar como controlo negativo, sendo constituído pela *Master Mix*, sem qualquer tipo de ADN.

3. Preparação da *Master Mix*:

<i>Buffer</i>	12,5 μ L X N
<i>Primer</i>	10,5 μ L X N

4. Distribuir 23 μ L da *Master Mix* por cada poço;
5. Nos primeiros 16 poços coloca-se as diluições de *ADN Standard* realizadas anteriormente. No poço NTC coloca-se apenas os 23 μ L da *Master Mix*. Nos poços seguintes colocam-se 2 μ L de ADN extraído das amostras pretendidas. Os últimos dois poços são colocados dois controlos:
9947 – Controlo feminino
007 – Controlo masculino.

Anexo III – Protocolo utilizado para Amplificação de ADN por PCR com o QIAGEN® Multiplex PCR Kit

1. Colocar os tubos *ependorf* previamente identificados juntamente com um tudo *ependorf* que vai ser utilizado para fazer a mistura na câmara de fluxo laminar e ligar os raios U.V. para descontaminação;
2. Num lugar separado, proceder ao descongelamento dos reagentes do QIAGEN® *Multiplex PCR Kit*, dos *primers* e das respetivas amostras, caso seja necessário;
3. Preparação da *Master Mix*:

<i>Multiplex</i>	5 μ L X N
<i>PrimerMix</i>	1 μ L X N
Água sem RNase	3,5 μ L X N

4. Agitar no *vórtex* a alta velocidade, entre 4 a 5 segundos;
5. Colocar em cada tubo 9,5 μ L da *Master Mix* e adicionar 0,5 μ L de produto extraído correspondente a cada amostra, perfazendo um volume final de 10 μ L.

Ao controlo negativo é adicionada a mesma quantidade de Master Mix;
6. As amostras devem ser colocadas num termociclador de modo a serem amplificadas:

Temperatura	Duração
95°C	15 Minutos
94°C	30 Segundos
60°C	90 Segundos
72°C	60 Segundos
72°C	10 Minutos
4°C	+ ∞

O restante da amostra extraída deve ser armazenada a -20°C se for reutilizada a longo prazo ou a 4°C se a curto prazo.

Em todo o processo de preparação das amostras para amplificação é aconselhado o uso de placas de gelo de modo a que os reagentes, *primers* e a *Master Mix* não sofram oscilações de temperatura.

- No final da reação de amplificação, deve-se colocar em cada *ependorf* 1 µL das enzimas *ExoSAP-IT*® e 1 µL de amostra amplificada por PCR.

A *ExoSAP-IT*® encontra-se no congelador a -20°C, quando se retira para utilização deve ser mantida em gelo.

É novamente colocado no termociclador para se proceder ao tratamento enzimático.

Temperatura	Duração
37°C	45 Minutos
85°C	15 Minutos

Anexo IV – Protocolo utilizado para Minisequenciação com SNaPShot™ Multiplex Kit

1. Colocar num local seguro os reagentes do *SNaPShot™ Multiplex Kit* e os *SBE Primers* a descongelar;
2. Preparar a *Master Mix*:

<i>SNaPShot™ Multiplex Mix</i>	1 µL X N
<i>SBE Primers</i>	1,27 µL X N
Água sem RNase	1,23 µL X N

3. Agitar no *vórtex* a alta velocidade e repartir 3,5 µL da mistura por cada tubo que contém a amostra amplificada e purificada;
4. Colocar as amostras no termociclador;

Temperatura	Duração	
96°C	10 Segundos	25 Ciclos
50°C	5 Segundos	
60°C	30 Segundos	
4°C	+ ∞	

5. Após terminada a reação de minisequenciação deve-se adicionar 2 μ L de *SAP*[®] (*Roche*[®]) a cada amostra e colocar novamente no termociclador.

A enzima *SAP*[®] (*Roche*[®]) encontra-se no congelador a -20°C, quando se retira para utilização deve ser mantida em gelo, assim como a *Master Mix* durante a distribuição pelos poços.

Temperatura	Duração
37°C	60 Minutos
85°C	15 Minutos

Anexo V – Protocolo utilizado para Eletroforese Capilar no sequenciador automático *ABI Prism® 3130 GeneticAnalyzer*

1. Preparar a *Master Mix*, já previamente descongelada num sítio seguro:

<i>Hi-Di™</i> Formamida	9 µL X N
<i>GeneScan® 120 LIZ Size Standard</i>	0,5 µL X N

Esta mistura é composta por *Hi-Di™* Formamida, uma solução desnaturante que tem como função quebrar as pontes de hidrogénio entre as cadeias complementares dos produtos resultantes da PCR.

O outro componente da mistura é o *GeneScan® 120 LIZ size standard*, um padrão interno que permite a atribuição de tamanhos. Este *size* foi especificamente desenhado para ser utilizado com o *SNaPSHOT™ Multiplex Kit*.

2. Agitar, a baixa velocidade, no *vórtex* durante 5 segundos;
3. Distribuir 9,5 µL pelos poços numa placa de 96 poços;
4. Adicionar 0,5 µL do produto SBE nos respetivos poços de cada amostra;
5. Proceder a um pequeno *spin*, de modo a que a amostras na placa não contenham bolhas;
6. Desnaturar no termociclador durante 5 minutos a 95°C e de seguida, colocar no sequenciador automático.

Anexo VI – Tabela de Resultados

Tabela 14 - Caracterização dos haplogrupos do Cromossoma Y na população caucasiana, mais especificamente portuguesa, com o estudo de 9 Y-SNPs conforme descrito por *Brionet al.* (2004). A análise foi feita por comparação com a Árvore de Haplogrupos Binários do Cromossoma Y, proposta pelo YCC.

Y-SNP	SRY1532	M213	M9	M70	M22	Tat	92R7	M173	P25	Haplogrupo
Deteção	A - G	A - G	C - G	T - G	A - G	A - G	TC - T	A - C	C - AC	
Mutação	A → G	T → C	C → G	A → C	A → G	T → C	G → A	A → C	C → A	
1	G	C	G	A	A	T	A	C	C	R1*
2	G	C	G	A	A	T	A	C	C	R1*
3	G	C	G	A	A	T	A	C	C	R1*
4	G	C	G	A	A	T	A	C	C	R1*
5	G	T	C	A	A	T	A	C	C	BCD e E
6	G	T	C	A	A	T	A	C	C	BCD e E
7	G	C	C	A	A	T	A	C	C	F (xK)
8	G	C	C	A	A	T	A	C	C	F (xK)
9	G	C	G	A	A	T	A	C	C	R1*
10	G	C	C	A	A	T	A	C	C	F (xK)
11	G	C	G	A	A	T	A	C	C	R1*
12	G	C	C	A	A	T	A	C	C	F (xK)
13	G	C	G	A	A	T	A	C	C	R1*
14	G	C	C	A	A	T	A	C	C	F (XK)
15	G	C	C	A	A	T	A	C	C	F (xK)
16	G	C	G	A	A	T	A	C	C	R1*
17	G	C	C	A	A	T	A	C	C	F (xK)
18	G	C	C	A	A	T	A	C	C	F (xK)
19	G	C	G	A	A	T	A	C	C	R1*
20	G	C	G	A	A	T	A	C	C	R1*

Tabela 14 - Continuação

21	G	C	C	A	A	T	A	C	C	F (xK)
22	G	C	G	A	A	T	A	C	C	R1*
23	G	C	C	A	A	T	A	C	C	F (xK)
24	G	C	G	A	A	T	A	C	C	R1*
25	G	C	C	A	A	T	A	C	C	F (xK)
26	G	C	G	A	A	T	A	C	C	R1*
27	G	C	G	A	A	T	A	C	C	R1*
28	G	C	G	A	A	T	A	C	C	R1*
29	G	C	G	A	A	T	A	C	C	R1*
30	G	C	G	A	A	T	A	C	C	R1*
31	G	C	G	A	A	T	A	C	C	R1*
32	G	T	C	A	A	T	A	C	C	BCD e E
33	G	C	C	A	A	T	A	C	C	F (xK)
34	G	C	G	A	A	T	A	C	C	R1*
35	G	C	C	A	A	T	A	C	C	F (xK)
36	G	C	G	A	A	T	A	C	C	R1*
37	G	C	G	A	A	T	A	C	C	R1*
38	G	C	G	A	A	T	A	C	C	R1*
39	G	C	G	A	A	T	A	C	C	R1*
40	G	C	G	A	A	T	A	C	C	R1*
41	G	C	G	A	A	T	A	C	C	R1*
42	G	C	G	A	A	T	A	C	C	R1*
43	G	C	C	A	A	T	A	C	C	F (xK)
44	G	C	G	A	A	T	A	C	C	R1*
45	G	C	C	A	A	T	A	C	C	F (xK)
46	G	C	G	A	A	T	A	C	C	R1*
47	G	C	G	A	A	T	A	C	C	R1*
48	G	C	G	A	A	T	A	C	C	R1*

Tabela 14 - Continuação

49	G	C	G	A	A	T	A	C	C	R1*
50	G	C	G	A	A	T	A	C	C	R1*
51	G	C	G	A	A	T	A	C	C	R1*
52	G	T	C	A	A	T	A	C	C	BCD e E
53	G	C	G	A	A	T	A	C	C	R1*
54	G	T	C	A	A	T	A	C	C	BCD e E
55	G	C	G	A	A	T	A	C	C	R1*
56	G	C	G	A	A	T	A	C	C	R1*
57	G	C	G	A	A	T	A	C	C	R1*
58	G	C	G	A	A	T	A	C	C	R1*
59	G	C	G	A	A	T	A	C	C	R1*
60	G	C	G	A	A	T	A	C	C	R1*
61	G	C	C	A	A	T	A	C	C	F (xK)
62	G	C	G	A	A	T	A	C	C	R1*
63	G	C	G	A	A	T	A	C	C	R1*
64	G	C	G	A	A	T	A	C	C	R1*
65	G	C	G	A	A	T	A	C	C	R1*
66	G	C	G	A	A	T	A	C	C	R1*
67	G	C	G	A	A	T	A	C	C	R1*
68	G	T	C	A	A	T	A	C	C	BCD e E
69	G	C	G	A	A	T	A	C	C	R1*
70	G	C	G	A	A	T	A	C	C	R1*
71	G	C	G	A	A	T	A	C	C	R1*
72	G	C	G	A	A	T	A	C	C	R1*
73	G	C	G	A	A	T	A	C	C	R1*
74	G	C	C	A	A	T	A	C	C	F (xK)
75	G	C	G	A	A	T	A	C	C	R1*

Tabela 15 – Caracterização dos haplogrupos do Cromossoma Y na população cabo-verdiana, com o estudo de 9 Y-SNPs conforme descrito por *Brión et al.* (2004). A análise foi feita por comparação com a Árvore de Haplogrupos Binários do Cromossoma Y, proposta pelo YCC.

Y-SNP	SRY1532	M213	M9	M70	M22	Tat	92R7	M173	P25	Haplogrupo
Deteção	A - G	A - G	C - G	T - G	A - G	A - G	TC - T	A - C	C - AC	
Mutação	A → G	T → C	C → G	A → C	A → G	T → C	G → A	A → C	C → A	
1	G	C	G	A	A	T	A	C	C	R1 *
2	G	C	G	A	A	T	A	C	C	R1 *
3	G	C	G	A	A	T	A	C	C	R1 *
4	G	C	G	A	A	T	A	C	C	R1 *
5	G	T	C	A	A	T	A	C	C	BCD e E
6	G	T	C	A	A	T	G	C	C	BCD e E
7	G	C	G	A	A	T	A	C	C	R1 *
8	G	C	C	A	A	T	G	C	C	F (xK)
9	G	T	C	A	A	T	A	C	C	BCD e E
10	G	C	C	A	A	T	A	C	C	F (xK)
11	G	C	G	A	A	T	A	C	C	R1 *
12	G	C	C	A	A	T	A	C	C	F (xK)
13	G	C	G	A	A	T	A	C	C	R1 *
14	G	C	G	A	A	T	A	C	C	R1 *
15	G	C	G	A	A	T	A	C	C	R1 *
16	G	C	G	A	A	T	A	C	C	R1 *
17	G	C	G	A	A	T	A	C	C	R1 *
18	G	C	G	A	A	T	A	C	C	R1 *
19	G	C	G	A	A	T	A	C	C	R1 *
20	G	C	G	A	A	T	A	C	C	R1 *
21	G	C	C	A	A	T	G	C	C	F (xK)
22	G	C	G	A	A	T	A	C	C	R1 *

Tabela 15 - Continuação

23	G	C	G	A	A	T	A	C	C	R1 *
24	G	C	G	A	A	T	A	C	C	R1 *
25	G	C	G	A	A	T	A	C	C	R1 *
26	G	T	C	A	A	T	A	C	C	BCD e E
27	G	T	C	A	A	T	A	C	C	BCD e E
28	G	C	G	A	A	T	A	C	C	R1 *
29	G	C	G	A	A	T	A	C	C	R1 *
30	G	C	G	A	A	T	A	C	C	R1 *
31	G	T	C	A	A	T	G	A	C	BCD e E
32	G	C	G	A	A	T	A	C	C	R1 *
33	G	C	G	A	A	T	A	C	C	R1 *
34	G	C	G	A	A	T	A	C	C	R1 *
35	G	C	G	A	A	T	A	C	C	R1 *
36	G	C	G	A	A	T	A	C	C	R1 *
37	G	C	G	A	A	T	A	C	C	R1 *
38	G	C	G	A	A	T	A	C	C	R1 *
39	G	C	G	A	A	T	A	C	C	R1 *
40	G	C	G	A	A	T	A	C	C	R1 *
41	G	C	C	A	A	T	A	C	C	F (xK)
42	G	T	C	A	A	T	A	A	C	BCD e E
43	G	C	G	A	A	T	A	C	C	R1 *
44	G	T	C	A	A	T	A	C	C	BCD e E
45	G	T	C	A	A	T	A	C	C	BCD e E
46	G	C	G	A	A	T	A	C	C	R1 *
47	G	C	C	A	A	T	A	C	C	F (xK)
48	G	T	C	A	A	T	A	C	C	BCD e E
49	G	C	G	A	A	T	A	C	C	R1 *

Tabela 15 - Continuação

50	G	C	G	A	A	T	A	C	C	R1 *
51	G	C	C	A	A	T	A	C	C	F (xK)
52	G	T	C	A	A	T	A	C	C	BCD e E
53	G	C	G	A	A	T	A	C	C	R1 *
54	G	C	G	A	A	T	A	C	C	R1 *
55	G	C	C	A	A	T	G	A	C	F (xK)
56	G	T	C	A	A	T	A	C	C	BCD e E
57	G	C	G	A	A	T	A	C	C	R1 *
58	G	T	C	A	A	T	A	C	C	BCD e E
59	G	T	C	A	A	T	A	C	C	BCD e E
60	G	T	C	A	A	T	A	C	C	BCD e E
61	G	C	G	A	A	T	A	C	C	R1 *
62	G	C	G	A	A	T	A	C	C	R1 *
63	G	C	C	A	A	T	A	C	C	F (xK)
64	G	C	G	A	A	T	A	C	C	R1 *
65	G	C	G	A	A	T	A	C	C	R1 *
66	G	C	G	A	A	T	A	C	C	R1 *
67	G	C	G	A	A	T	A	C	C	R1 *
68	G	T	G	A	A	T	A	C	C	BCD e E
69	G	C	G	A	A	T	A	C	C	R1 *
70	G	C	G	A	A	T	A	C	C	R1 *
71	G	C	G	A	A	T	A	C	C	R1 *
72	G	C	G	A	A	T	A	C	C	R1 *
73	G	T	C	A	A	T	A	C	C	BCD e E
74	G	C	G	A	A	T	A	C	C	R1 *
75	G	T	C	A	A	T	A	C	C	BCD e E

ÍNDICE DE FIGURAS

Figura 1 - Representação da constituição de uma molécula de ADN, com o açúcar, o grupo fosfato e a base azotada. O nucleótido tem a letra que representa a base azotada presente. Adaptado de: http://www.mas-que-ciencia.com/category/genetica/genetica-molecular	14
Figura 2 - Cromossoma Y e os seus constituintes. Adaptado de Gusmão e Alves, 2010.	18
Figura 3 - Representação de um SNP. A sequência difere apenas num único par de bases. O polimorfismo é de C para T. Adaptado de Pinheiro, 2010.	22
Figura 4 – Esquema acerca da importância da aplicação dos SNPs na Genética Forense. Adaptado de Pinheiro, 2010.	23
Figura 5 - Árvore dos haplogrupos binários do cromossoma Y proposta pelo Y Chromosome Consortium em 2002 (The Y Chromosome Consortium, 2002). Atualmente contém cerca de 311 haplogrupos, reunindo 600 polimorfismos binários. Adaptado de Y Chromosome Consortium, 2002.	31
Figura 6 - Mapa de Cabo Verde, onde são visíveis as ilhas do Barlavento e as do Sotavento. Adaptado de: http://unesdoc.unesco.org/images/0013/001330/133020por.pdf	32
Figura 7 - Rota do tráfico de escravos. Partiam de vários pontos da costa africana com um ponto de paragem – Cabo Verde - daí seguiam para vários pontos do continente americano. Adaptado de: https://tracingafricanroots.wordpress.com/maps/upper-guinea	34
Figura 8 - Mapa de Portugal, com representação dos arquipélagos Açores e Madeira. Adaptado de: http://www.mapadeportugal.net/indicadistritos.asp	37

Figura 9 - Representação esquemática da reação de PCR. Adaptado de Butler, 2005.	48
Figura 10 - Representação de ciclos de PCR. É dividido em 3 diferentes temperaturas que são repetidas entre 25 a 35 ciclos. A 94°C o ADN é desnaturado, como efeito há a separação da cadeia dupla em duas cadeias simples, em que cada uma das cadeias irá servir de molde para cópia. A 60°C sucede-se o emparelhamento dos primers com a cadeia molde de ADN. Aos 72°C a DNA Polimerase permite a extensão dos primers e a formação de uma nova cadeia. Adaptado de Pinheiro, 2008.	48
Figura 11 - Esquema de purificação com ExoSAP-IT® (usb®). Adaptado de: tudogeral.sib.uc.pt.	52
Figura 12 - Esquema geral de funcionamento da ExoSAP- IT®. Adaptado de: http://labs.icb.ufmg.br/lbem/protocolos/exosap.html	53
Figura 13 - Esquema geral da reação de minisequenciação com o SNaPSHOT™ Kit. Adaptado de Applied Biosystems, 2012.	54
Figura 14 - Representação do funcionamento da Reação de Minisequenciação com o SNaPSHOT™ Multiplex Kit. Adaptado de Applied Biosystems, 2012.....	60
Figura 15 - Representação de um eletroforegrama com os 9 Y-SNPs em estudo...	61
Figura 16 - Representação dos fragmentos do padrão interno GeneScan® 120 LIZ size standard.	62
Figura 17 - Árvore dos haplogrupos binários do cromossoma Y proposta pelo Y Chromosome Consortium. Adaptado de Y Chromosome Consortium, 2002.	65

Figura 18 - Árvore dos haplogrupos binários do cromossoma Y proposta pelo YCC e revista por Karafet et al. (2008). Adaptado de Karafet et al, 2008.....	66
Figura 19 - Eletroforegrama dos 9 Y-SNPs que se apresentam sob a forma de fragmentos com diferentes tamanhos e cores, conforme a sua variação nucleotídica.	67
Figura 20 - Eletroforegrama característico do haplogrupo R1*.....	68
Figura 21 - Eletroforegrama característico do haplogrupo F (xK).	68
Figura 22 - Eletroforegrama característico dos haplogrupos B,C,D ou E.	69
Figura 23 - Representação gráfica dos haplogrupos da população caucasiana.....	71
Figura 24 - Representação gráfica dos haplogrupos da população cabo-verdiana .	71

ÍNDICE DE TABELAS

Tabela 1 – Comparação das diferentes características dos marcadores genéticos entre SNPs e STRs.	27
Tabela 2 – Resumo dos métodos laboratoriais para a análise dos SNPs do Cromossoma Y.....	45
Tabela 3 - Seleção de Y-SNPs distribuídos no multiplex, com informação da mutação e respetiva deteção, segundo Brión et al. (2004).	46
Tabela 4 - Composição do kit comercial QIAGEN® Multiplex PCR Kit.	49
Tabela 5 - Caraterização dos primers utilizados para amplificação dos Y-SNPs.....	50
Tabela 6 - Programa para amplificação das amostras.....	51
Tabela 7 - Representação dos fluorocromos e o seu respetivo ddNTP assim como a cor que é visível na eletroforese. Em relação ao ddCTP, para uma melhor observação, a cor que surge no eletroforegrama é preto no lugar do amarelo.	55
Tabela 8 - Caraterização dos primers para a reação de minisequenciação dos Y-SNPs.	56
Tabela 9 - Composição do kit comercial SNaPSHOT™ Multiplex Kit.....	57
Tabela 10 - Programa de PCR para a Minisequenciação.	58
Tabela 11 - Programa descrito no termociclador para tratamento enzimático com a SAP® (Roche®).	59
Tabela 12 - Frequências dos haplogrupos presentes na população caucasiana, especificamente portuguesa.....	70
Tabela 13 - Frequências dos haplogrupos presentes na população cabo-verdiana	70
Tabela 14 - Caracterização dos haplogrupos do Cromossoma Y na população caucasiana, mais especificamente portuguesa, com o estudo de 9 Y-SNPs conforme	

descrito por Brión et al. (2004). A análise foi feita por comparação com a Árvore de Haplogrupos Binários do Cromossoma Y, proposta pelo YCC..... **99**

Tabela 15 - Caracterização dos haplogrupos do Cromossoma Y na população cabo-verdiana, com o estudo de 9 Y-SNPs conforme descrito por Brión et al. (2004). A análise foi feita por comparação com a Árvore de Haplogrupos Binários do Cromossoma Y, proposta pelo YCC..... **102**

ÍNDICE DE ABREVIATURAS

ADN – Ácido desoxirribonucleico

ADNnu – Ácido desoxirribonucleico nuclear

AIMs – *Ancestry Informative Markers* – Marcadores informativos de ancestralidade

CEE – Comunidade Económica Europeia

ddNTP – dideoxinucleótidos trifosfatados

dNTP – desoxinucleótidos trifosfatos

DIP – *Delection /Insertion Polymorphism*

Exo-SAP – *Exonuclease I e Shrimp Phosphatase Alkaline*

F_{st} – Índice de fixação

HLA – *Human Leukocyte Antigen*

InDel – Inserção/Deleção

INE – Instituto Nacional de Estatística

INMLCF, IP. – Instituto Nacional de Medicina Legal e Ciências Forenses, Instituto Público

IPC – *Internal Positive Control*

Mb – Milhões de pares de bases

MSY – *Male specific Y region*

NRY – *Non recombining region of the Y chromosome*

PAR 1 – *Pseudoautosomal Region 1*

PAR 2 – *Pseudoautosomal Region 2*

PCR – *Polymerase Chain Reaction* – Reação de polimerase em cadeia

pb – pares de bases

rpm – rotações por minuto

RT-PCR – *real time – Polymerase Chain Reaction*

SAP – *Shrimp Phosphatase Alkaline*

SBE – *Single Base Extension*

SNP – *Single Nucleotide Polymorphism*

SRY – *Sex determination region of the Y chromosome*

STR – *Short Tandem Repeat*

YCC – *Y Chromosome Consortium*

Yp – braço curto do cromossoma Y

Yq – braço longo do cromossoma Y

Y-SNP – *Single Nucleotide Polymorphism of the Y chromosome.*